

MAGISTERUPPSATS I BIBLIOTEKS- OCH INFORMATIONSVETENSKAP  
VID BIBLIOTEKS- OCH INFORMATIONSVETENSKAP/BIBLIOTEKSHÖGSKOLAN  
2006:27  
ISSN 1404-0891

# En utvärdering av två pearl growing-metoder i ISI Web of Science

HELENA AHLIN  
GUNILLA PETTERSSON



HÖGSKOLAN I BORÅS

© **Helena Ahlin/Gunilla Pettersson**  
Mångfaldigande och spridande av innehållet i denna uppsats  
– helt eller delvis – är förbjudet utan medgivande.

Svensk titel: En utvärdering av två pearl growing-metoder i ISI Web of Science

Engelsk titel: An Evaluation of Two Pearl Growing Methods in ISI Web of Science

Författare: Helena Ahlin, Gunilla Pettersson

Kollegium: 2

Färdigställt: 2006

Handledare: Per Ahlgren

Abstract: The purpose of this thesis is to evaluate two pearl growing methods in ISI Web of Science. The methods tested are Find Related Records (FRR), an automatic service that uses bibliographic coupling, and Author Keywords (AK), where searches are performed using the two first keywords from this field. Both search methods were used on twenty initial documents and the first twenty hits were evaluated for relevance using a binary scale (0,1). The relevance criteria were based on specific information needs connected to every initial document. The measures used to evaluate the effectiveness of the search methods were Precision at DCV=10, (P(10)), and Uninterpolated Average Precision (AP). Wilcoxon's signed rank test was used to evaluate if there were any significant differences between the two methods. The results showed that FRR achieved the highest average values for both P(10) and AP. However, neither of the methods excelled in effectiveness, something that the Wilcoxon's test corroborated. When it comes to FRR's results one could state that bibliographic coupling does not always perform well when a specific information need is requested. One reason for AK's results is that keywords can be too wide or too narrow and consequently affect the retrieval of information. If a search starts with a relevant initial document, pearl growing can be a useful search method, especially if both FRR and AK are combined.

Nyckelord: Pearl growing, ISI Web of Science, sökmetoder, utvärdering, återvinningseffektivitet, IR

## Innehållsförteckning

1. Inledning .....	1
1.1 Syfte och frågeställningar .....	2
1.2 Disposition.....	3
2. Teoretisk bakgrund .....	4
2.1 IR .....	4
2.1.1 IR-modeller.....	4
2.2 Utvärdering av IR-system.....	5
2.2.1 Cranfield .....	7
2.2.2 TREC .....	7
2.2.3 Effektivitetsmått .....	8
2.3 Relevansbegreppet.....	9
2.4 Sökprocessen .....	10
2.4.1 Olika typer av sökningar.....	10
2.4.2 Pearl growing.....	11
2.5 Bibliografisk koppling .....	12
3. Tidigare forskning .....	13
4. Metod .....	16
4.1 ISI Web of Science .....	16
4.2 Initialdokumenten.....	18
4.3 Tillvägagångssätt .....	19
4.3.1 Relevansbedömning.....	19
4.3.2 Effektivitetsberäkningar .....	20
4.3.3 Relevansantaganden med avseende på strukna dokument .....	22
5. Resultat .....	23
5.1 Återvinnings effektivitet.....	23
5.2 Signifikanstestning .....	26
5.3 Bästa/Sämsta tänkbara utfall.....	28
6. Diskussion.....	32
6.1 Relevansbedömning.....	32
6.2 Resultatet .....	33
6.2.1 Precision (10).....	33
6.2.2 Uninterpolated Average Precision.....	34
6.2.3 Pearl growing.....	36
6.3 Utfallet av signifikanstestningen .....	37
6.4 Bästa/Sämsta tänkbara utfall.....	37
6.4.1 Precision (10).....	38
6.4.2 Uninterpolated Average Precision.....	38
6.5 Koppling till tidigare forskning .....	39
7. Slutsatser .....	43
8. Sammanfattning .....	45
Referenser .....	47
Bilaga 1. Initialdokumenten.....	49

# 1. Inledning

Inom bibliotekssektorn, och särskilt på forskningsbibliotek, använder man dagligen databaser av olika slag i sökandet efter vetenskapligt material. Studenter och forskare är ofta i behov av sådant material, många gånger i form av artiklar. Inom vissa ämnen, särskilt där utvecklingen går snabbt, är artiklar ett bra sätt att hålla sig ajour på. För återvinning av dessa anses databaser vara pålitliga informationskällor. Att känna till och behärska databaser och deras sökmöjligheter är viktiga kunskaper som innehas av många bibliotekarier. Det är viktigt att dessa kunskaper hålls uppdaterade för att kunna guida användarna och på så sätt kunna erbjuda bra och snabb hjälp när detta behövs.

Det finns olika typer av databaser. Vissa är ämnesspecifika, till exempel LISA och ERIC, som specialiserar sig inom ämnena biblioteks- och informationsvetenskap (B&I), respektive pedagogik, psykologi och utbildning. Andra är ämnesövergripande, såsom Academic Search Elite som är en fulltextdatabas, och ISI Web of Science som är en portal för databaserna Science Citation Index Expanded, Social Sciences Citation Index och Arts & Humanities Citation Index. ISI står för The Institute for Scientific Information och ägs numera av The Thomson Corporation. Den här uppsatsens empiriska studie kommer att utföras i ISI Web of Science.

Återvinningsmetoderna skiljer sig databaser emellan. Gränssnitten är formade på olika sätt, är mer eller mindre komplicerade, och kan upplevas som svåra av ovana användare. Det är av stor vikt att användaren tar sig tid att lära sig hur databasens gränssnitt fungerar och tar reda på vilka sökmöjligheter som finns. Ett exempel på en av de mest använda sökstrategierna är building blocks där termer i facetter kombineras med hjälp av booleska operatorer till olika söksträngar.<sup>1</sup> Det här är en av många sökstrategier. Ofta behöver man dock kombinera olika för att uppnå bäst resultat.

Ännu ett exempel på en sökstrategi är pearl growing, den strategi som den här uppsatsen fokuserar på. Genom pearl growing utgår man ifrån ett relevant dokument, en pärla eller initialdokument, för att återvinna liknande dokument, alltså dokument inom samma ämnesområde. Den klassiska metoden att utföra pearl growing på är att plocka termer från det relevanta initialdokumentets titel, abstrakt och/eller brödtext och använda termerna i nästa sökning. En annan metod som kan ses som en typ av pearl growing är bibliografisk koppling, en metod som förenar ett initialdokument med andra dokument som har minst en gemensam referens med initialdokumentet. ISI Web of Science har implementerat bibliografisk koppling som en del av sitt utbud.

Den här uppsatsen kommer att fokusera på hur två metoder av pearl growing kan tillämpas. Genom att testa metoderna i ISI Web of Science kan för- och nackdelar med att utgå från ett initialdokument jämföras och utvärderas. Intentionen med den här granskningen är att se om pearl growing är en effektiv och meningsfull metod, framförallt vid sökning efter vetenskapligt material inom ett specifikt ämnesområde.

---

<sup>1</sup> Harter, Stephen P. (1986). *Online Information Retrieval: Concepts, Principles, and Techniques*, s. 172.

Då sökningar i databaser är en integrerad del av arbetet på ett bibliotek är det intressant att se vilka sökmetoder som uppbringar bäst resultat. Skälet till att ISI Web of Science är passande att studera i det här sammanhanget är tvåfaldigt. Dels är ISI Web of Science:s databaser ämnesövergripande, varpå den lämpar sig för till exempel ämnet biblioteks- och informationsvetenskap som har en tvärvetenskaplig karaktär. Dels har ISI Web of Science två hjälpmedel som gagnar pearl growing-metoden: Find Related Records och Author Keywords.

*Find Related Records (FRR)* finns i den bibliografiska posten och är en funktion som kopplar ihop dokument som delar minst en gemensam referens. Genom att klicka på ikonerna för Find Related Records presenteras en träfflista på dokument som delar minst en referens med initialdokumentet. Dokumenten rankas efter hur många gemensamma referenser de har med initialdokumentet, en så kallad kopplingsstyrka. Find Related Records-metoden kallas, som tidigare nämnts, för bibliografisk koppling.

*Author Keywords (AK)* är författarens nyckelord till det egna dokumentet. Nyckelorden presenteras i ett eget fält i den bibliografiska posten. Den här uppsatsen använder sig av sökningar på de två första termerna i detta fält som en andra pearl growing-metod. En sökfråga konstrueras med hjälp av de två första termerna som binds samman med AND-operatoren. På så sätt förväntas sökningen generera dokument inom samma ämnesområde som initialdokumentet.

De här två pearl growing-metoderna kommer i uppsatsen att jämföras vid sökandet efter dokument inom ämnet biblioteks- och informationsvetenskap. Det intressanta med en jämförelse är att sökmetoden där AK används baseras på individers subjektiva val av nyckelord. FRR däremot är en automatisk tjänst. Om den automatiska tjänsten visar sig fungera på ett tillfredställande sätt skulle dess användbarhet kunna bekräftas och tjänsten särskilt rekommenderas åt bibliotekarier och kanske även ersätta tidskrävande manuella pearl growing-metoder.

AK-metoden är idag inte en automatisk tjänst men man måste dock tänka på att den skulle kunna implementeras som en sådan. Därför är det viktigt att veta hur bra den presterar och hur effektiv den är i jämförelse med FRR.

## **1.1 Syfte och frågeställningar**

Syftet med den här uppsatsen är att testa och jämföra två pearl growing-metoder för att se vilken utav dem som uppvisar bäst återvinningseffektivitet vad gäller specifika informationsbehov inom ämnet biblioteks- och informationsvetenskap.

Uppsatsen utgår från följande tre frågeställningar:

- Vilken precision uppvisar de olika metoderna?
- Vad beror eventuella skillnader i precision på?
- På vilket sätt är pearl growing en effektiv och meningsfull metod, framförallt vid sökning efter vetenskapligt material inom ett specifikt ämnesområde?

## 1.2 Disposition

Kapitel 2 kommer att behandla uppsatsens teoretiska bakgrund. En kort presentation av fältet *Information Retrieval* (IR) kommer att ges, där bland annat de tre klassiska IR-modellerna introduceras. I kapitel 2 kommer även relevansbegreppet att problematiseras och en bakgrund till pearl growing och bibliografisk koppling att ges. Kapitel 3 behandlar tidigare forskning inom ämnesområdet bibliografisk koppling. Då ingen specifik forskning hittats om pearl growing har detta utelämnats. Kapitel 4 kommer att redovisa för uppsatsens metod. Här ingår även en presentation av databasportalen ISI Web of Science. I kapitel 5 redovisas den empiriska studiens resultat som sedan diskuteras i kapitel 6. Kapitel 7 innehåller uppsatsens slutsatser. Slutligen kommer kapitel 8 att sammanfatta uppsatsen.

## 2. Teoretisk bakgrund

### 2.1 IR

Termen Information Retrieval (IR), informationsåtervinning, började användas under 1950-talet.<sup>2</sup> Inom IR studeras representation, förvaring, organisation och tillgängliggörande av information. Representationen och organisationen av information ska underlätta för användaren att hitta rätt information.<sup>3</sup> Ett IR-systems funktion är inte att ge direkta svar på användarens informationsbehov utan att hänvisa till dokument som gör detta. Ett IR-system kan tillhandahålla bibliografiska poster, det vill säga detaljer rörande det sökta dokumentet, såsom till exempel titel, ISBN-nummer och abstrakt. I många fall kan IR-systemet även innehålla dokumentet i full-text.<sup>4</sup>

Användarens informationsbehov måste omvandlas från naturligt språk till en sökfråga (engelsk term: *query*) för att kunna användas i ett IR-system.<sup>5</sup> Sedan gäller det för systemet att matcha sökfrågan med databasens dokument. Användarstudier har visat att användare tar med sig olika sorters bakgrund, kompetens, och erfarenheter när de använder ett IR-system. Man bör tänka på att ett IR-system används av både barn och vuxna, användare med olika mängder tid och tålmod, i privata såsom offentliga miljöer samt för olika informationsbehov.<sup>6</sup> Syftet med ett IR-system är att finna rätt sorts information till rätt användare.<sup>7</sup> Användaren måste själv avgöra vilken information som är relevant för informationsbehovet, men IR-systemet bistår genom att "tolka" innehållet i dokumenten och ranka dem efter relevansgrad i förhållande till sökfrågan. Relevansbegreppet är centralt inom IR. IR-systemets yttersta syfte är att återvinna alla relevanta dokument till en sökfråga och så få icke-relevanta dokument som möjligt.<sup>8</sup>

Det vanligaste sättet att representera dokument i större samlingar är att tilldela varje dokument ett antal indexeringstermer och skapa en inverterad fil för databasen. Den inverterade filen innehåller indexeringstermer med pekare till varje dokument som indexerats med de här termerna. Tekniken medföljs ofta av en algoritm som tillåter användaren att använda booleska kombinationer av termer som sökfrågor och kombinera träffar på ett lämpligt och automatiskt sätt.<sup>9</sup>

#### 2.1.1 IR-modeller

I ett IR-system bestäms relevansen av ett dokument utifrån likheten mellan en sökfråga och dokumenten i databasen. Olika matematiska modeller har konstruerats för detta ändamål.<sup>10</sup> Här följer en kort presentation av de tre vanligaste modellerna.

---

<sup>2</sup> Chowdhury, G. G. (1999). *Introduction to Modern Information Retrieval*, s. 1.

<sup>3</sup> Baeza-Yates, Ricardo & Ribeiro-Neto, Berthier (1999). *Modern Information Retrieval*, s. 1.

<sup>4</sup> Chowdhury (1999), s. 1-2.

<sup>5</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 1.

<sup>6</sup> Marchionini, Gary (1995). *Information Seeking in Electronic Environments*, s. 98.

<sup>7</sup> Chowdhury (1999), s. 2.

<sup>8</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 2.

<sup>9</sup> Marchionini (1995), s. 24.

<sup>10</sup> Chowdhury (1999), s. 160.

## Booleska modellen

Den booleska modellen bygger på kombinationen av söktermer med hjälp av de så kallade booleska operatörens, AND, OR och NOT. Modellen kan upplevas som enkel men har sina nackdelar. Dess återvinningsstrategi baseras på en binär skala, antingen anses ett dokument vara relevant eller så anses det vara icke-relevant. Någon partiell matchning förekommer alltså inte. Man kan därför säga att den booleska modellen mer är en dataåtervinningsmodell än en informationsåtervinningsmodell. En annan nackdel är att det ibland kan vara problematiskt att översätta ett informationsbehov till ett booleskt uttryck.<sup>11</sup> Det finns risk att sökfrågan blir för bred eller för precis. Blir den för bred kan antalet återvunna dokument bli oöverskådligt, vilket försvårar för användaren att hitta relevant information. På samma sätt kan en alltför precis sökfråga generera alldeles för få dokument och göra så att användaren missar relevant information.<sup>12</sup>

## Vektormodellen

Vektormodellen erbjuder partiell matchning. Det här uppnås genom att tilldela icke-binära vikter till termer i sökfrågor och dokument. Vikterna används för att beräkna likheten mellan sökfråga och dokument. Dokumenten sorteras sedan i fallande ordning utifrån grad av likhet, varpå man i en träfflista även kan se dokument som enbart delvis matchar dokumentet. Det anses att vektormodellen ger en mer precis rankning av dokument än vad booleska modellen gör. På så sätt matchar träfflistan användarens informationsbehov bättre.<sup>13</sup>

## Probabilistiska modellen

Den probabilistiska modellen utgår från sannolikheten att ett dokument är relevant givet en viss sökfråga. I en dokumentsamling finns det vissa dokument som är helt relevanta för användarens informationsbehov. För att återvinna dem måste rätt indexeringstermer användas, termer som användaren inte känner till på förhand. Först måste en gissning utföras, där termer som antas vara rätt används. Användaren får inspektera träfflistan och bestämma vilka dokument som är relevanta och vilka som inte är det. Oftast behöver användaren enbart gå igenom de första dokumenten på listan för att kunna avgöra detta. Utifrån den första träfflistans resultat förfinas sökningen och körs igen. Den här processen upprepas ett antal gånger varpå träfflistan antas förbättras för varje gång med avseende på att innehålla relevanta dokument.<sup>14</sup>

## 2.2 Utvärdering av IR-system

Utvärdering av IR-system görs för att förvissa sig om dess prestanda och värde. Inom IR kan man testa hur bra två system presterar i förhållande till varandra för att se vilket som är bäst eller hur bra ett system presterar för att sedan eventuellt förbättra det. För att

---

<sup>11</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 26.

<sup>12</sup> Chowdhury (1999), s. 161.

<sup>13</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 27.

<sup>14</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 31.



göra detta finns enligt G.G. Chowdhury två olika mätvariabler, *effectiveness* och *efficiency*. Effectiveness mäter hur bra systemet når upp till sin målsättning, vilket kan innebära hur bra systemet klarar av att ta fram relevanta dokument och samtidigt undvika de icke-relevanta. Efficiency mäter hur systemet når upp till sin målsättning ekonomiskt sett och kan till exempel mätas genom sådana faktorer som vilken den minsta kostnaden är som systemet klarar för att arbeta effektivt. De här kostnaderna räknas på hur snabbt systemet svarar, vilken tid det behövs för att sätta sig in i systemet och så vidare.<sup>15</sup>

Det finns två huvudsakliga faktorer för att evaluera ett IR-system, prestanda och kostnad. Prestanda involverar begreppet relevans som är ett mått på sambandet mellan dokument och användare.<sup>16</sup> Mer om relevansbegreppet i kapitel 2.3.

C. W. Cleverdon kom 1966 fram till sex kriterier för utvärdering av återvinningssystem:

- systemets förmåga att frambringa alla relevanta dokument, recall
- systemets förmåga att endast frambringa relevanta dokument, precision
- tiden det tar att generera ett svar på en sökfråga
- den intellektuella och fysiska ansträngning som krävs av användaren
- sökresultatets utformning
- samlingens täckning<sup>17</sup>

Den här uppsatsen utvärderar två sökmetoder av pearl growing i ISI Web of Science. En av metoderna, FRR, är automatisk och tillhör IR-systemet. Genom att utvärdera den metoden så utvärderar man också en del av systemet. Även om det inte är själva IR-systemet som utvärderas i uppsatsen så är det ändå av intresse att titta närmare på hur kriterier för sådana utvärderingar har uppkommit.

Anledningen till att en presentation av Cranfield-projektet och TREC-konferensen tas med i denna teoretiska bakgrund beror på att Cranfield-projektet utvecklade effektivitetsmåten precision och recall som ingår i den här uppsatsen och TREC-konferensen, å sin sida, använder sig av pooling-metoden, vilket också är av intresse i den här uppsatsen.

Slutligen kan det nämnas att utvärderingsstudier kan göras ur två synpunkter, antingen från ett managementorienterat perspektiv eller från ett användarperspektiv. Många forskare inom IR förespråkar användarperspektivet.<sup>18</sup> Det är då intressant att se om informationssystemet hjälper användaren att tillfredställa sitt informationsbehov.<sup>19</sup> Den här uppsatsen kan användas i ett användarperspektiv eftersom dess resultat i slutändan kan leda till mer kunskap om sökmetoderna samt tankar om hur och varför bibliotekarier och andra användare kan nyttja dem på bästa sätt.

---

<sup>15</sup> Chowdhury (1999), s. 200.

<sup>16</sup> Chowdhury (1999), s. 202.

<sup>17</sup> Chowdhury (1999), s. 203.

<sup>18</sup> Chowdhury (1999), s. 201.

<sup>19</sup> Kagolovsky, Yuri & Moehr, Jochen R. (2003). Current Status of the Evaluation of Information Retrieval. *Journal of Medical Systems*, vol. 27, no. 5, s. 410.

### 2.2.1 Cranfield

De första betydande utvärderingsstudierna startades 1957 i Cranfield, Storbritannien, under ledning av Cleverdon och gav en ny dimension till forskningen kring utvärdering av IR-system. Under Cranfield-projekten, Cranfield 1 och 2, koncentrerade man sig främst på utvärdering av indexeringsystem, både manuella och automatiska. Cranfields största bidrag var dock att de utvecklade lämpliga metoder för utvärdering av återvinning. Dessa studier ledde även fram till de mått, precision och recall (definieras i avsnitt 2.2.3), som ofta används för att utvärdera IR-system samt deras inverterade relation till varandra.<sup>20</sup>

Cranfield-projekten etablerade en standard för experiment inom utvärdering av informationsåtervinning. Studierna ledde till intensiva diskussioner, framförallt kring studiens relevansbedömningar och begreppet relevans. Diskussioner kring svårigheterna med begreppet relevans pågår inom IR-studier än idag. Cranfield-projekten har även blivit kritiserade för att vara för småskaliga, detta är dock något som inte kan sägas om TREC-konferensen.<sup>21</sup>

### 2.2.2 TREC

Konferensen TREC (Text REtrieval Conference) startades 1992 som en del av TIPSTER Text program och är sponsrat av the National Institute of Standards and Technology (NIST) samt USA:s försvarsdepartement.<sup>22</sup> Konferensen hålls årligen och testkollektionen TREC räknas idag som referens för informationsåtervinning, detta på grund av dess väldiga storlek (över en miljon dokument) samt grundliga experimenterande.<sup>23</sup>

Syftet med konferensen är att stödja forskning inom IR genom att tillhandahålla den infrastruktur som är nödvändig för storskaliga utvärderingar av textåtervinningsmetoder. TREC:s mål är att gynna IR-forskning baserad på stora testkollektioner. TREC vill även öka kommunikationen av forskningsidéer genom ett öppet forum mellan industri, den akademiska världen och staten. Ett mål är också att snabba på överföringen av teknologi från forskningslabb till kommersiella produkter genom att demonstrera betydande förbättringar inom återvinningsmetoder på verkliga problem. Slutligen vill man förbättra tillgängligheten av utvärderingstekniker för den akademiska världen och industrin samt utveckla metoder som är lämpligare för dagens system.<sup>24</sup>

För varje konferens designas en uppsättning referensexperiment som de medverkande forskningsgrupperna får använda för att jämföra sina återvinningsystem. TREC-kollektionen består av tre delar: dokumenten, topics (exempel på informationsbehov) samt en uppsättning relevanta dokument för varje informationsbehov.<sup>25</sup>

---

<sup>20</sup> Chowdhury (1999), s. 215-218.

<sup>21</sup> Kagolovsky (2003), s. 415-416.

<sup>22</sup> Text Retrieval Conference (TREC) (2004-01-30). Overview. <http://trec.nist.gov/overview.html> [2006-02-17]

<sup>23</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 84-85.

<sup>24</sup> Text Retrieval Conference (TREC) (2004-01-30). [2006-02-17].

<sup>25</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 85-86.

På konferensen skapas en pool av relevanta dokument, detta görs genom att man tar de topprankade dokumenten för ett topic från de olika deltagarna och samlar i en pool. Dokumenten där relevansbedöms sedan av människor. Tekniken för att få fram dokument för relevansbedömning kallas pooling-metoden och bygger på två antaganden, dels att merparten av de relevanta dokumenten finns i poolen, dels att dokumenten utanför poolen kan betraktas som icke-relevanta.<sup>26</sup>

Konferensen innehåller huvudsakligen två återvinningsuppgifter: den ena uppgiften kallas ad hoc och innebär att en serie nya informationsbehov körs mot en fast dokumentdatabas. Den andra är en routing-uppgift där en serie fasta informationsbehov körs mot en föränderlig dokumentdatabas.<sup>27</sup> TREC försöker därmed efterlikna verkliga förhållanden. Ett exempel på ad hoc är bibliotekskataloger och ett exempel på routing är självuppdaterande nyhetstjänster.

### 2.2.3 Effektivitetsmått

För att räkna ut precision och recall krävs det att man delar in dokumenten i fyra olika typer: relevanta och återvunna dokument, icke-relevanta och återvunna dokument, relevanta och ej återvunna dokument samt icke-relevanta och icke återvunna dokument. Recall är proportionen återvunna relevanta dokument av det totala antalet relevanta dokument i en dokumentsamling. Precision är proportionen av återvunna dokument som är relevanta för sökfrågan.<sup>28</sup>

$$\text{Recall} = \frac{\text{antalet återvunna relevanta dokument}}{\text{antalet relevanta dokument i samlingen}}$$

$$\text{Precision} = \frac{\text{antalet återvunna relevanta dokument}}{\text{antalet återvunna dokument}}$$

Under Cranfield-experimenten kom man fram till att precision och recall har ett inverterat förhållande till varandra. Försöker man öka den ena så minskar den andra, högre recall kan endast åstadkommas på bekostnad av precisionen och vice versa.<sup>29</sup> Det finns dock vissa forskare som menar på att det här förhållandet inte alltid stämmer.<sup>30</sup>

Recall kan vara svårt att räkna ut i stora dokumentsamlingar då man oftast inte är medveten om hur många relevanta dokument som finns. Det kan vara svårt att räkna ut ett exakt värde på recall. För det här krävs en samling med känt antal relevanta dokument för en sökfråga. I många fall skulle även ett mått som kombinerar recall och precision vara bra, eftersom de är relaterade mått men fångar olika aspekter av uppsättningen återvunna dokument.<sup>31</sup> Måtten precision och recall har blivit mycket

---

<sup>26</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 89.

<sup>27</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 89.

<sup>28</sup> Kagolovsky (2003), s. 418.

<sup>29</sup> Large, Andrew, Tedd, Lucy A. & Hartley R. J. (2001). *Information Seeking in the Online Age : Principles and Practice*, s. 281.

<sup>30</sup> Se Chowdhury (1999), s. 207 eller Large, Tedd & Hartley (2001), s. 281.

<sup>31</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 81.

kritiserade för att de bygger på relevansbedömningar. Trots att det skapats många andra mått för att bedöma IR-system så används dock fortfarande de relevansbaserade måtten precision och recall mest.<sup>32</sup>

## 2.3 Relevansbegreppet

Som tidigare nämnts är relevansbegreppet centralt inom IR. Termen har dock gett upphov till en hel del diskussioner eftersom det förekommer en mängd olika definitioner av relevans. Stefano Mizzaro tar upp olika exempel på termer som används för begreppet: ”‘relevance’, ‘topicality’, ‘utility’, ‘usefulness’, ‘system relevance’, ‘user relevance’ [...]”. Ibland används termerna synonymt, ibland inte.<sup>33</sup>

Relevans anses allmänt vara relationen mellan å ena sidan informationsresurser och å andra sidan representationer av användarnas informationsbehov. Mizzaro beskriver de tre enheterna inom informationsresurser såsom: det fysiska *dokumentet*, *dokument-surrogatet*, till exempel en bibliografisk post, samt *information*, det användaren får/tar emot vid läsandet av ett dokument.<sup>34</sup>

Relevansbedömningen av dokument grundar sig i ett initialt informationsbehov hos användaren. Mizzaro har fyra begrepp kopplade till användarens informationsbehov. Först uppstår ett informationsbehov, något som Mizzaro kallar för *Real Information Need (RIN)*, ett behov som sedan blir *Perceived Information Need (PIN)*, en mental representation av informationsbehovet. Användaren uttrycker PIN genom en fråga, en representation av PIN i naturligt språk (*request*), och slutligen formaliseras frågan genom en sökfråga (*query*), en representation av frågan genom till exempel booleskt språk. Det här förloppet kan komprimeras till tre operationer: perception, formulering och formalisering, operationer som dock inte är helt självklara.<sup>35</sup>

Perceptionsfasen från RIN till PIN kan bli problematisk eftersom användaren söker efter något okänt vilket ibland kan innebära att användaren inte vet vad hon/han vill veta. Formuleringssfasen kompliceras av svårigheterna med att uttrycka ett informationsbehov i ord och den slutliga formaliseringsfasen är svår eftersom systemspråket kan upplevas som svårt av användaren.<sup>36</sup>

Relevansbegreppet är inte bara komplicerat på grund av alla definitioner som finns och på den ej självklara gången från informationsbehov till konkret sökfråga. Tidpunkten när ett dokument relevansgranskas har också en stor betydelse. Ett dokument kan anses vara relevant vid en viss tidsperiod för att anses vara icke-relevant vid ett senare tillfälle, och tvärtom. Andra faktorer som spelar in vid relevansgranskning är komplexiteten av användarens intresseområde, vad informationen ska användas till samt kontexten inom vilken information söks.<sup>37</sup>

---

<sup>32</sup> Kagolovsky (2003), s. 418-419.

<sup>33</sup> Mizzaro, Stefano (1998). How Many Relevances in Information Retrieval?. *Interacting with Computers*, vol. 10, no. 3, s. 305.

<sup>34</sup> Mizzaro (1998), s. 305.

<sup>35</sup> Mizzaro (1998), s. 306.

<sup>36</sup> Mizzaro (1998), s. 307.

<sup>37</sup> Mizzaro (1998), s. 308-310.

Som redan nämnts finns det en uppsjö av olika definitioner av relevansbegreppet. Man skiljer till exempel på *utility* och *relevance*: ”a document can be relevant to the query, but its utility to the user will be zero.”<sup>38</sup> Även om ett dokument bedöms som relevant av systemet kan dokumentet alltså vara icke-relevant för användaren.<sup>39</sup> IR-system kan sägas ha en objektiv relevanstolkning eftersom de antar att en sökfråga verkligen motsvarar en användares informationsbehov och att relevansbedömningar kan utföras av vilken användare som helst. Den här utgångspunkten baseras på en binär relevansbedömning, relevant eller icke-relevant.<sup>40</sup> En sådan här relevansbedömning görs på en binär skala (0,1), där 0 står för icke-relevant och 1 står för relevant.<sup>41</sup> Det finns de som tycker att en relevansgranskning utifrån en binär skala är ett sätt att förenkla relevansbedömningen eftersom dokument ofta är mer eller mindre relevanta.<sup>42</sup>

En binär relevansskala har använts i den här uppsatsen. Anledningen till att just en sådan relevansskala använts, trots dess begränsningar, är att det i praktiken ofta inte finns tid eller möjlighet att göra någon noggrannare bedömning av dokument med till exempel en flergradig relevansskala. I den här uppsatsen utgår relevanskriterierna från fastställda informationsbehov, vilka har skapats utifrån initialdokument och de ämnen som dessa behandlar. De dokument som i den här uppsatsen ansetts vara relevanta har uppfyllt relevanskriterierna.

Sammanfattningsvis kan det konstateras att relevansbegreppet är relativt och att en relevansbedömning beror på bedömarens förkunskaper. Vetskapen om att man kan se på relevansbegreppet på olika sätt är dock det som gör det hela till ett problem.<sup>43</sup>

## 2.4 Sökprocessen

Informationssökningsprocessen börjar med ett informationsbehov. Gary Marchionini skiljer på begreppen *information retrieval* och *information seeking*. *Retrieval* förutsätter enligt Marchionini att det man söker efter är ”känt”. *Seeking* kopplas ihop med förvärvning av kunskap, en problemorienterad process där lösningar eventuellt kan hittas. *Retrieval* kopplas ihop med databaser, *seeking* med besvarandet av frågor och inläring.<sup>44</sup>

### 2.4.1 Olika typer av sökningar

Andrew Large et al. menar att sökningar kan delas in i tre breda kategorier: *known-item-sökning*, *faktasökning* och *ämnessökning*. Vid *known-item-sökning* försöker användaren

---

<sup>38</sup> Cooper, i Kagolovsky (2003), s. 419.

<sup>39</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 1-2.

<sup>40</sup> Kagolovsky (2003), s. 418.

<sup>41</sup> Baeza-Yates & Ribeiro-Neto (1999), s. 26.

<sup>42</sup> Efthimiadis, Efthimis N. (1996). Query Expansion. Ingår i Williams, Martha E. ed. *Annual Review of Information Science and Technology: Vol. 31*. s. 137-138.

<sup>43</sup> Saracevic, Tefko (1975). Relevance: A Review of and a Framework for the Thinking on the Notion in Information Science. *Journal of the American Society for Information Science*, vol. 26, s. 325.

<sup>44</sup> Marchionini (1995), s. 5-6.

lokalisera information om något som är känt på förhand. Ett exempel kan vara sökningar i en bibliografisk databas eller en bibliotekskatalog. Här har användaren en känd utgångspunkt, till exempel namnet på en författare.<sup>45</sup>

Vid faktasökning vill man fastställa eller bekräfta en specifik faktafråga. Det finns en specifik fråga och man vill få ett svar på den. Vid sådana sökningar är inte sökstrategin det viktigaste utan snarare att hitta en lämplig informationskälla, sökningen i sig är ofta rätt enkel. Internets sökmotorer är väl använda källor i dessa lägen trots risken att få fram stora mängder icke-relevant information i träfflistorna.<sup>46</sup>

Ämnessökningar skiljer sig från de två föregående kategorierna. Known-item-sökning och faktasökning anses vara avklarade så fort det sökta objektet återvunnits eller ett lämpligt svar erhållits. Ämnessökningar är däremot inte lika förbehållslösa, eftersom de försöker hitta information om ett visst ämne. Vanligtvis vet användaren inte, vid sökningens början, hur mycket information som kan tänkas finnas om ämnet i fråga. Svårt, om inte omöjligt, att veta är också *när* all relevant information återfunnits. Ämnessökningar kräver större sökkunskaper av användaren. Här krävs en genomtänkt sökstrategi.<sup>47</sup> I den här uppsatsen ligger fokus på ämnessökningar.

Marchionini skiljer mellan två sorters sökningsstrategier, analytiska strategier och browsing-strategier.<sup>48</sup> Analytiska strategier är målorienterade och systematiska och kräver en noggrann planering. Browsing, å andra sidan, baseras på igenkännande av relevant information.<sup>49</sup> Browsing-strategier används av slutanvändaren själv men analytiska strategier kan användas av till exempel en bibliotekarie, eller annan intermediär, som hjälper användaren. I praktiken, skriver Marchionini, används dock oftast en mix av dessa båda strategier. Människor tenderar att använda sig av browsing i hög utsträckning men analytiska strategier är mer effektiva när det gäller stora dokumentsamlingar.<sup>50</sup>

## 2.4.2 Pearl growing

Pearl growing-metoden utgår från ett specifikt initialdokument som användaren vet är relevant för informationsbehovet, en så kallad pärla. I klassisk pearl growing använder man sig av nyckelord från initialdokumentet med vilka användaren fortsätter att söka och återvinna fler relevanta dokument.<sup>51</sup> Även om pearl growing börjar med ett väldefinierat och relevant initialdokument kräver den här sökstrategin större interaktion mellan användaren och systemet än vad andra sökstrategier gör.<sup>52</sup> När användaren erhållit tillräckligt med information eller när inga nya söktermer kan hittas är sökningen klar.<sup>53</sup> Det kan dock vara svårt för användaren att avgöra när ”growing”-processen ska

---

<sup>45</sup> Large, Tedd & Hartley (2001), s. 35-36.

<sup>46</sup> Large, Tedd & Hartley (2001), s. 35-36.

<sup>47</sup> Large, Tedd & Hartley (2001), s. 35-36.

<sup>48</sup> Marchionini (1995), s. 8.

<sup>49</sup> Harter (1986), s. 76.

<sup>50</sup> Marchionini (1995), s. 8.

<sup>51</sup> Marchionini, (1995), s. 78.

<sup>52</sup> Marchionini, (1995), s. 79.

<sup>53</sup> Harter (1986), s. 184.

avslutas, det vill säga *när* användningen av nyckelord ska avslutas, något som kräver ett stort engagemang från användarens sida.<sup>54</sup>

I den här uppsatsen motsvarar AK-metoden klassisk pearl growing medan FRR-metoden kan ses som en variant av pearl growing, en variant som involverar bibliografisk koppling.

## 2.5 Bibliografisk koppling

”Bibliometri innebär kvantitativa analyser, främst av vetenskaplig litteratur, där målet är att beskriva vetenskapens och teknologins utveckling. Bibliometrins grundantagande är att litteraturen avspeglar forskningsaktiviteten”. Inom bibliometrin forskas det följaktligen i vetenskaplig kommunikation.<sup>55</sup>

Bibliografisk koppling används inom bibliometrin och är ett mått på semantisk likhet mellan dokument. Måttet tillämpas på par av dokument och man kan säga att två dokument  $d_1$  och  $d_2$  är *bibliografiskt kopplade* om de har minst en gemensam referens. Termen introducerades 1963 av M.M. Kessler från Massachusetts Institute of Technology (MIT).<sup>56</sup>

Bibliografisk koppling är en typ av källanalys, där publikationers referenser granskas och jämförs med varandra. En grupp vetenskapliga publikationer anses ha en inbördes relation om de har minst en gemensam referens. Bibliografisk koppling används för att gruppera publikationer inom ett ämnesområde och därmed visa på nätverk som beskriver publikationernas förhållanden sinsemellan.<sup>57</sup>

Ett problem med källanalyser är att de inte tar hänsyn till att en publikation kan inkluderas i en annan publikations källförteckning av olika orsaker och att alla källciteringar därför inte kan ses som likvärdiga.<sup>58</sup> Forskare citerar till exempel inte alltid alla de dokument som haft betydelse för deras arbete. Vissa titlar utelämnas till exempel ibland på grund av att de anses tillhöra disciplinens vedertagna allmänbildning.<sup>59</sup>

---

<sup>54</sup> Marchionini (1995), s. 78-79.

<sup>55</sup> Kärki, Riita & Kortelainen, Terttu (1998). *Introduktion till bibliometri*, s. 9-10.

<sup>56</sup> Kessler, M. M. (1963). Bibliographic Coupling Between Scientific Papers. *American Documentation*, 14, s. 10-11.

<sup>57</sup> Kärki & Kortelainen (1998), s. 22.

<sup>58</sup> Kärki & Kortelainen (1998), s. 22.

<sup>59</sup> Kärki & Kortelainen (1998), s. 92.

### 3. Tidigare forskning

Följande kapitel behandlar tidigare forskning inom ämnesområdet bibliografisk koppling. Då ingen specifik forskning hittats om pearl growing har detta utelämnats.

**Eugene Garfield** introducerade 1955 Science Citation Index (SCI),<sup>60</sup> en citeringsdatabas som idag kan nås via ISI Web of Science i en "Expanded"-version. Därmed ökade intresset för att undersöka relationen mellan publicerade dokument efter de citeringar de innehåller.<sup>61</sup> Kopplat till detta finner man bibliografisk koppling, ett begrepp skapat av **M. M. Kessler** som säger att en referens som delas av två dokument definieras som en kopplingsenhet dem emellan. Kopplingen anses bli starkare ju fler gemensamma referenser dokumenten delar.<sup>62</sup> En av hans undersökningar gjordes på 8521 artiklar ur tidskriften *Physical Review*. Kessler har utifrån kopplingsenheter definierat två kriterier A och B:

Kriterium A: Ett antal dokument utgör en relaterad grupp  $G_A$  om varje dokument i gruppen har åtminstone en kopplingsenhet till ett givet testdokument  $P_o$ . Kopplingsstyrkan mellan  $P_o$  och något dokument ur  $G_A$  mäts genom antalet kopplingsenheter ( $n$ ) dem emellan.  $G_A^n$  är proportionen av  $G_A$  som är länkad till  $P_o$  genom  $n$  kopplingsenheter.<sup>63</sup>

Kriterium B: Ett antal dokument utgör en relaterad grupp  $G_B$  om varje dokument i gruppen har åtminstone en kopplingsenhet till alla andra dokument i gruppen.<sup>64</sup>

För den här uppsatsen är kriterium A relevant, då  $G_A$  motsvarar dokumenten i databasen som har minst en kopplingsenhet till  $P_o$ , ett initialdokument.

Kessler sammanfattar den bibliografiska kopplingens egenskaper som återvinningsverktyg på följande sätt:

- Bibliografisk koppling är oberoende av ord och språk. Därmed undviks språkliga svårigheter som syntax och språkvanor.
- Det behövs ingen expertbedömning av dokumenten.
- Bibliografisk koppling sträcker sig till det förflutna lika mycket som till framtiden.
- Metoden producerar inte en statisk klassifikation för en given vetenskaplig artikel. Den genomgår förändringar som reflekterar den aktuella användningen av det vetenskapliga ämnesfältet.
- Artiklar som delar en grupp av kopplingar med en given artikel kan betraktas som dess logiska referens och ses som ett substitut för dess egna referenser.  $G_A$  kan ses som  $P_o$ :s logiska referenser.<sup>65</sup>

---

<sup>60</sup> Ramer, Sheryl L. (2005). Site-ation Pearl Growing: Methods and Librarianship History and Theory. *Journal of the Medical Library Association*, vol. 93, no. 3, s. 399.

<sup>61</sup> Martyn, John (1964). Bibliographic Coupling. *Journal of Documentation*, vol. 20, s. 236.

<sup>62</sup> Martyn (1964), s. 236.

<sup>63</sup> Kessler (1963), s. 10.

<sup>64</sup> Kessler (1963), s. 10.



Kessler har förblivit ett viktigt namn inom bibliografisk koppling men hans forskning har även granskats kritiskt av andra. Här följer ett antal undersökningar/utvärderingar som gjorts genom åren.

**John Martyn** visade 1964 viss skepsis vad gäller Kesslers teori och menade på att bibliografisk koppling endast är en indikation på att det existerar en trolig relation mellan två dokument, en koppling av okänt värde.<sup>66</sup>

**Bella Hass Weinberg** recenserade de teoretiska och praktiska tillämpningar som gjorts inom bibliografisk koppling fram till 1974 då hennes artikel publicerades. Hon visade även på användbarheten av bibliografisk koppling som återvinningsverktyg inom IR.<sup>67</sup>

Weinberg kom fram till att bibliografisk koppling kan vara ett användbart verktyg för att studera "the 'science of science'", alltså citeringsmönster, de mest citerade tidskrifterna och så vidare.<sup>68</sup> En invändning mot Kesslers arbete var att han håller sig till ett specifikt ämnesområde, där dokumenten redan antas ha en koppling. För att visa om bibliografisk koppling fungerar i en komplex, tvärvetenskaplig miljö bör dock ett test i större skala göras, till exempel i SCI, enligt Weinberg.<sup>69</sup>

**Subir K Sen** och **Shymal K Gan** utvecklade 1983 idén om bibliografisk koppling samt formaliserade den matematiskt. De fokuserade på kopplingsstyrka, det vill säga att ju fler gemensamma referenser som delas mellan dokument desto större är chansen att dokumenten har någonting gemensamt. De föreslog ett mått på kopplingsstyrka, Coupling Angle, med avseende på klustring.<sup>70</sup>

**George Vladutz** och **James Cook** tittade 1984 närmare på fenomenet bibliografisk koppling i SCI och kom fram till att tekniken är ett effektivt sätt att återvinna ämnesrelaterade dokument på. Vladutz och Cook baserade sin slutsats på ett experiment där 10 000 tidskriftsartiklar slumpmässigt valdes ut ur SCI år 1981 och användes för att undersöka om det förekom någon koppling mellan dem och stora delar av SCI:s databas (ca 100 000 dokument valdes bort då de inte innehöll referenslistor).<sup>71</sup>

Experimentet prövade Kesslers teori om den ökade graden av koppling beroende på hur många gemensamma referenser som delas av dokumenten. En utvärdering genomfördes också för att undersöka ämnesrelevansen mellan de dokument som kopplas ihop. Resultatet blev att 90 % av artiklarna producerade en grupp av två eller fler kopplade dokument.<sup>72</sup> I utvärderingen om ämnesrelevans visade det sig att 86 % av de starkast

---

<sup>65</sup> Kessler (1963), s. 11.

<sup>66</sup> Martyn (1964), s. 236.

<sup>67</sup> Weinberg, Bella Hass (1974). Bibliographic Coupling: a Review. *Information Storage and Retrieval*, vol. 10, no. 5/6, s. 189.

<sup>68</sup> Weinberg (1974), s. 195.

<sup>69</sup> Weinberg (1974), s. 195.

<sup>70</sup> Sen, Subir K. & Gan, Shymal K. (1983). A Mathematical Extension of the Idea of Bibliographic Coupling and its Applications. *Annals of Library Science and Documentation*, vol. 30, no. 2, s. 78-82.

<sup>71</sup> Vladutz, George & Cook, James (1984). Bibliographic Coupling and Subject Relatedness. *Proceedings of the 47<sup>th</sup> ASIS Annual Meeting*, vol. 21, s. 204.

<sup>72</sup> Vladutz & Cook (1984), s. 204.

kopplade dokumenten bedömdes vara "well related by subject."<sup>73</sup> Vladutz och Cook kom till slutsatsen: "It is hypothesized that the bibliographic coupling techniques may prove to be the easiest approximation to an algorithm for revealing the semantically closest neighbors of publications."<sup>74</sup>

**H.P.F. Peters, R.R. Braam och A.F.J. van Raan** visade med sin forskning att dokument med en bibliografisk koppling har en betydligt större ämnesrelation än andra dokument. Deras forskning visar också på att dokument som citeras flitigt oftast citeras inom det egna ämnesfältet.<sup>75</sup> Forskningen baseras på "word-profile similarity analysis".<sup>76</sup> Denna likhet (similarity) är större inom en grupp dokument som delar en referens till ett ofta citerat dokument än bland en grupp dokument som saknar denna koppling.<sup>77</sup>

**Bo Jarneving** kom 2005 ut med en avhandling med titeln *The Combined Application of Bibliographic Coupling and the Complete Link Cluster Method in Bibliometric Science Mapping* där han presenterar en metod för kartläggning av forskningsstrukturer, så kallat science mapping. Jarneving föreslår en kombination av bibliografisk koppling och complete link clustering, två metoder som kompletterar varandra. Genom bibliografisk koppling påvisas likhet mellan dokument, en kognitiv koppling, och genom complete link clustering grupperas dokument. I complete link clustering har varje medlem i klustret en koppling till alla andra medlemmar i klustret.<sup>78</sup>

---

<sup>73</sup> Vladutz & Cook (1984), s. 204 och 206.

<sup>74</sup> Vladutz & Cook (1984), s. 206.

<sup>75</sup> Peters, H.P.F., Braam, R.R. och van Raan, A.F.J. (1995). Cognitive Resemblance and Citation Relations in Chemical Engineering Publications. *Journal of the American Society for Information Science*, vol. 46, no. 1, s. 10, 19 och 21.

<sup>76</sup> Peters, Braam & van Raan (1995), s. 19.

<sup>77</sup> Peters, Braam & van Raan (1995), s. 21.

<sup>78</sup> Jarneving, Bo (2005). *The Combined Application of Bibliographic Coupling and the Complete Link Cluster Method in Bibliometric Science Mapping*, s. 30.

## 4. Metod

### 4.1 ISI Web of Science

Den empiriska studiens experimentella miljö utgjordes av ISI Web of Science, en betaldataportal. The Institute for Scientific Information (ISI) har skapat tre citeringsindex. Science Citation Index, som numera har tillägget Expanded, grundades 1963 och innehåller referenser till naturvetenskaplig, medicinsk och teknisk litteratur. Social Sciences Citation Index (1972) bevakar samhällsvetenskaplig litteratur och Arts & Humanities Citation Index (1980) humanistisk litteratur.<sup>79</sup> Citeringsindexen skapades från början för informationsåtervinning och det speciella med dem är att de innehåller dokumentens referenslistor, något som många andra bibliografiska databaser saknar.<sup>80</sup>

1997 introducerade ISI databasportalen ISI Web of Science som gjorde de tre citeringsdatabaserna mer tillgängliga för användarna med sitt användarvänliga gränssnitt.<sup>81</sup> ISI Web of Science tillhandahåller, enligt egen utsago, "information from approximately 8,700 of the most prestigious, high impact research journals in the world."<sup>82</sup>

ISI Web of Science erbjuder bland annat Quick Search, General Search och Advanced Search. Den här empiriska studiens tester utfördes i General Search. Oberoende av sökmetod går det att välja att söka i alla databaserna samtidigt eller i var och en för sig. I General Search är det möjligt att söka på fält såsom Topic, Author, Group Author, Source Title, Publication Year och Address. Det finns möjlighet att specificera vilket språk och/eller dokumenttyp man vill återvinna. Vidare kan man även begränsa sig tidsmässigt från 1986 och framåt. Det finns olika möjligheter att sortera träfflistan som fås efter en sökning i General Search. För den här empiriska studien valdes den förinställda sorteringen efter senaste datum. Detta innebär att dokumenten sorteras efter det datum då publiceringen processades av Thomson ISI med den senast processade högst upp i träfflistan.<sup>83</sup>

Figur 1 illustrerar hur en fullständig bibliografisk post i ISI Web of Science kan se ut. Förutom gängse uppgifter såsom dokumentets titel, författare, källa, dokumenttyp, språk och abstrakt, innehåller den bibliografiska posten även en ikon för den automatiska tjänsten FRR samt ett fält med författarens egna nyckelord, det som i uppsatsen kallas för AK. Genom tjänsten FRR har ISI Web of Science implementerat bibliografisk koppling, närmare bestämt Kesslers kriterium A.<sup>84</sup>

I den bibliografiska posten finns också ett fält som heter *KeyWords Plus*, detta är ett fält som består av ofta förekommande ord och fraser från titlar i ett dokument referenser. Orden och fraserna i *KeyWords Plus* behöver dock inte nödvändigtvis förekomma i

---

<sup>79</sup> Kärki & Kortelainen (1998), s. 24.

<sup>80</sup> Kärki & Kortelainen (1998), s. 24.

<sup>81</sup> Tenopir, Carol (2001). The Power of Citation Searching. *Library Journal*, vol. 126, no. 18, s. 39-40.

<sup>82</sup> Web of Science – Thomson Scientific. <http://scientific.thomson.com/products/wos/> [2006-03-29].

<sup>83</sup> Web of Science – Thomson Scientific. [2006-03-29].

<sup>84</sup> Se kapitel 3.

dokumentets egen titel eller i AK-fältet. Länken *Cited References* bidrar med en lista på dokumentets referenser. Länken *Times Cited* visar de artiklar som citerat dokumentet.<sup>85</sup>

**Web of Science®**

WELCOME

**Error! Hyperlink reference not valid.**

GENERAL SEARCH CITED REF SEARCH SEARCH HISTORY ADVANCED SEARCH

Full Record

Record 1 of 1 (Set #1)

**Title:** Philosophical foundations and research relevance: issues for information research

**Author(s):** [Wilson TD](#)

**Source:** JOURNAL OF INFORMATION SCIENCE 29 (6): 445-452 2003

**Document Type:** Article

**Language:** English

**Cited References:** 32 **Times Cited:** 2 [FIND RELATED RECORDS](#)

**Error! Hyperlink reference not valid.**

**Abstract:** This paper examines three ideas that affect the nature of research in information science: (1) the fact of the incoherent nature of information science, which results from the concept of 'information' being dealt with at different integrative levels; (2) the lack of an over-arching philosophical framework that might guide the development of methods; and (3) the problem of grounding research in the reality of everyday professional practice. It is suggested that, at one level, phenomenology offers an integrative philosophical perspective that might also help to resolve the research/practice split.

**Author Keywords:** information science; librarianship; research; philosophical framework; methodologies; models; relevance; professional practice; knowledge organization; integrative levels; phenomenology

**KeyWords Plus:** SYSTEMS; MODELS; BEHAVIOR

**Addresses:** Wilson TD (reprint author), Univ Sheffield, Dept Informat Sci, Western Bank, Sheffield, S Yorkshire S10 2TN England  
Univ Sheffield, Dept Informat Sci, Sheffield, S Yorkshire S10 2TN England

**Publisher:** SAGE PUBLICATIONS LTD, 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND

**Subject Category:** COMPUTER SCIENCE, INFORMATION SYSTEMS; INFORMATION SCIENCE & LIBRARY SCIENCE

**IDS Number:** 757UQ

**ISSN:** 0165-5515

Record 1 of 1 (Set #1)

**Output This Record**

Bibliographic Fields

PRINT E-MAIL SAVE

EXPORT TO REFERENCE SOFTWARE

Or add it to the Marked List for later output and more options.

ADD TO MARKED LIST **Error!**

**Hyperlink reference not valid.**  
[0 articles marked]

**Create Citation Alert**

CREATE CITATION ALERT

Receive e-mail alerts on future citations to this record. (Requires registration.)

Holdings GO

**View record in**  
[Journal Citation Reports](#)

**Figur 1.** Fullständig bibliografisk post i ISI Web of Science.

<sup>85</sup> Se bibliografiska postens hjälp-funktion i ISI Web of Science.

## 4.2 Initialdokumenten

Initialdokumenten, 20 stycken, som användes i uppsatsen tillhör disciplinen biblioteks- och informationsvetenskap. I *Dictionary for Library an Information Science*, delas disciplinen in i Information Studies, Information Science och Library Science:

### **Information Studies**

An umbrella term used at some universities for a curricular division that includes library and information science (LIS) and allied fields (informatics, information management, etc.).

### **Information Science**

The systematic study and analysis of the sources, development, collection, organization, dissemination, evaluation, use, and management of information in all its forms, including the channels (formal and informal) and technology used in its communication.

### **Library Science**

The professional knowledge and skill with which recorded information is selected, acquired, organized, stored, maintained, retrieved, and disseminated to meet the needs of a specific clientele.<sup>86</sup>

För att finna initialdokumenten genomfördes sökningar i ISI Web of Science databaser Science Citation Index, Social Sciences Citation Index och Arts & Humanities Citation Index. Att sökningarna utfördes i alla tre databaserna beror på ämnet B&I:s tvärvetenskapliga karaktär. Dokument inom ämnet B&I finns inom vitt skilda ämnesområden såsom teknik, samhällsvetenskap och humaniora.

För att få inspiration till söktermer lästes en artikel ur *Journal of the American Society for Information Science (JASIS)* som ger en bred bild av disciplinen Informationsvetenskap (Information Science).<sup>87</sup> Artikeln presenterar de mest citerade och därmed mest inflytelserika författarna inom området samt ämnets olika delfält. Således gav artikeln tips på författarnamn att söka på. Förutom författarnamn söktes det även på termer som hör till ämnesområdet B&I, till exempel *classification theory*.

Följande begränsningar utfördes när sökningarna efter initialdokumenten genomfördes:

- Språk: English
- Dokumenttyp: Article
- Ålder på återvunna dokument: From 1995 to 2006.

Begränsningarna gjordes eftersom studenter och forskare antas behärska det engelska språket samt att en övervägande del av allt vetenskapligt material publiceras på engelska. Att artiklar valts som dokumenttyp beror på att forskning ofta redovisas i denna form. Slutligen så bestämdes tidsbegränsningen till en tioårsperiod för att få relativt nytt material.

---

<sup>86</sup>Reitz, Joan M. (2004). *Dictionary for Library and Information Science*.

<sup>87</sup>White, Howard D. & McCain, Katherine W. (1998). Visualizing a Discipline: An Author Co-Citation Analysis of Information Science, 1972-1995. *Journal of the American Society for Information Science*, vol. 49, no. 4, s. 327-355.

Initialdokumenten hade ett antal villkor att uppfylla. Den bibliografiska posten måste innehålla AK, vilket många poster inte hade. AK-fältet har dock bara funnits som fält i bibliografiska poster sedan 1991. I de fall AK ingick i den bibliografiska posten måste sökningen på de två första nyckelorden generera minst 20 dokument. På samma sätt måste funktionen FRR också generera minst 20 dokument. Så fort ett initialdokument uppfyllde dessa krav, valdes det ut till den empiriska studien. När 20 initialdokument hade återvunnits påbörjades den empiriska studien.

### **4.3 Tillvägagångssätt**

Den empiriska studien utgick från ovan nämnda initialdokument. Två stycken pearl growing-metoder testades på varje enskilt initialdokument. Den första metoden använde funktionen FRR som finns som ikon i den bibliografiska posten och som länkar vidare till dokument som har minst en gemensam referens med initialdokumentet. Eftersom det är en automatisk funktion så erbjuder den dock inte möjligheten att avgränsa efter tid, språk och dokumenttyp.

Andra metoden gick ut på att finna relevanta dokument genom att söka på författarnas egna nyckelord till initialdokumenten. De första två termerna i initialdokumentets AK-fält bands samman med den booleska operatoren AND och en sökning gjordes i General Search, i Topic-fältet. Anledningen till att just de två första nyckelorden valdes ut i den här uppsatsen var för att alla initialdokument skulle behandlas på lika villkor vid sökningarna. Dessutom ansågs två stycken nyckelord vara ett lämpligt antal, dels för att sökningen hade blivit för bred med en enda sökterm, dels för att sökningen blivit för smal med fler söktermer.

AK-metoden avgränsades även till att bara återvinna artiklar, detta för att forskare ofta är intresserade av sådana dokument. Dessutom hjälper en sådan avgränsning till att undvika att återvinna vissa dokumenttyper som finns med i ISI Web of Science som till exempel poesi, teaterrecensioner, brev och så vidare. Sökningarna avgränsades även till att bara återvinna dokument på engelska.

Document Cutoff Value (DCV) = 20, det vill säga att endast de 20 högst rankade dokumenten för varje sökning granskades.

#### **4.3.1 Relevansbedömning**

För att kunna relevansbedöma dokumenten konstruerades informationsbehov för varje initialdokument av uppsatsens författare. Informationsbehoven tillkom genom en läsning av initialdokumentens titel och abstrakt i den bibliografiska posten. För att inte gynna AK-metoden granskades inte nyckelorden i AK-fältet vid bestämmandet av informationsbehovet. Efter att metoderna FRR och AK använts, relevansbedömdes träfflistorna enligt en binär skala (0,1) där de bedömdes såsom icke-relevanta alternativt relevanta av uppsatsens författare.

Relevansbedömningen utgick från fastställda relevanskriterier kopplade till det informationsbehov som uppsatsförfattarna associerat med varje initialdokument.<sup>88</sup> Kriterierna färdigställdes innan sökningarna utfördes. Relevansbedömningen utfördes enskilt av uppsatsens författare och jämfördes sedan för att nå en så objektiv bedömning av dokumentens relevans som möjligt. Eventuellt olika bedömningar granskades gemensamt av författarna för att nå en slutgiltig relevansbedömning.

Vid sökning med hjälp av AK är det möjligt att initialdokumentet dyker upp i träfflistan på grund av att även initialdokumentet indexerats under dessa termer. Om initialdokumentet återfanns i träfflistorna ersattes det dokumentet med nästa dokument i listan. På så sätt granskades även det 21:a dokumentet. Samma sak gällde om en dubblett dök upp i listan, det vill säga om två olika tidskrifter hade publicerat exakt samma dokument. Dubbletter och initialdokument som återkom i träfflistan ingick alltså inte i den empiriska studien och det var alltså 20 dokument per träfflista som ingick i studien.

Om en bibliografisk post saknade abstrakt, vilket inte var ovanligt, söktes dokumentet upp på annat sätt då detta var möjligt. De här sökningarna genomfördes i olika databaser via Samsök, i LISA, ERIC samt i sökmotorerna Google Scholar och Google. Om abstrakt inte var tillgängligt online söktes dokumentet upp i pappersform i den mån det fanns att tillgå på högskolebiblioteket i Borås. Var det inte möjligt att komma över dokumentet ansågs dokumentet inte möjligt att relevansgranska och ersattes därmed med nästa dokument i listan.

Enbart dokument på engelska relevansgranskades i den här uppsatsen. Dokument på annat språk än engelska ersattes av näst följande dokument på engelska i träfflistan. AK-metoden gick att avgränsa till engelska men i FRR-metoden var det inte möjligt eftersom det är en automatisk tjänst där inga avgränsningar kan göras.

### 4.3.2 Effektivitetsberäkningar

Jämförelsen metoderna emellan skedde genom användningen av följande mått:

1. *Precision vid DCV=10*,  $P(10)$ . Antalet relevanta dokument i positionerna 1-10 dividerat med 10.
2. *Uninterpolated Average Precision*, (AP). Precisionen beräknas vid varje relevant dokument i listan av 20 återvunna dokument,  $DCV = 20$ . Sedan summeras precisionsvärdena och summan divideras med antalet kända relevanta dokument. AP-värdenas medelvärde heter *Mean Uninterpolated Average Precision* (MAP) och redovisas i resultatavsnittet.

---

<sup>88</sup> Se Bilaga 1. Initialdokumenten

Ovanstående mått illustreras med följande exempel:

P(10): Om en av metoderna, metod 1, genererar 3 relevanta dokument bland de 10 första i träfflistan blir precisionen  $3/10 = 0,3 = 30\%$ . Om metod 2 genererar 5 relevanta dokument bland de 10 första i träfflistan blir precisionen  $5/10 = 0,5 = 50\%$ .

AP: Antag att det finns 7 relevanta dokument som motsvarar informationsbehovet. Om en av metoderna, metod 1, genererar 5 relevanta dokument som finns på positionerna 1, 3, 7, 11, 12 och den andra metoden, metod 2, genererar 3 relevanta dokument på positionerna 2, 3, 9 räknas AP ut på följande sätt:

Först beräknas precision för varje relevant dokument i listan och precisionsvärdena summeras:

$$\text{Metod 1. } (1/1) + (2/3) + (3/7) + (4/11) + (5/12) = 2,88.$$

$$\text{Metod 2. } (1/2) + (2/3) + (3/9) = 1,50.$$

Antalet unika relevanta dokument i båda metodernas träfflistor slås ihop i en pool.<sup>89</sup> Om vi utgår ifrån att metoderna i exemplet tillsammans genererade 7 unika relevanta dokument, ett relevant dokument återfanns i båda träfflistorna, består poolen i det här fallet av 7 dokument. Metodernas sammanlagda precisionsvärden divideras med 7.

$$\text{Metod 1. } 2,88/7 = 0,41 = 41\%.$$

$$\text{Metod 2. } 1,50/7 = 0,21 = 21\%.$$

Metod 1 visar på ett AP-värde på 41 % och uppvisar alltså ett bättre resultat än metod 2 som får 21 % i det här exemplet.

Anledningen till att både mått 1 och 2 används i uppsatsen är för att de kompletterar varandra. P(10) visar på hur många relevanta dokument som finns bland de tio första platserna i träfflistan. Det är ofta så att användaren vare sig har tid eller lust att granska fler än de dokument som finns på träfflistans första sida, en sida som ofta innehåller tio dokument. Det är därför av stor vikt att så många relevanta dokument som möjligt finns bland de tio första, det vill säga att precisionen där är hög.

AP har en inbyggd recall-aspekt. Genom detta antas det att man vet vilka de relevanta dokumenten är, eftersom de ingår i en pool. På så sätt påminner det om en testkollektion där man vet vilka dokument som är relevanta för ett informationsbehov. Även här är det av stor vikt att sökmetoderna återvinner relevanta dokument med placeringar högt upp i listan. I AP påverkar de två olika sökmetoderna som jämförs varandra, eftersom deras sökresultat påverkar poolen och därmed resultatet.

---

<sup>89</sup> Se avsnittet om TREC, 2.2.2.



### 4.3.3 Relevansantaganden med avseende på strukna dokument

I de fall där abstrakt saknades i den bibliografiska posten, och det inte gick att hitta vare sig på elektronisk väg eller i pappersformat, så ersattes det dokumentet av näst följande dokument med abstrakt i träfflistan. Om de dokument i träfflistan som saknade abstrakt inte hade saknat det hade de gått att relevansbedöma och därför ingått i den empiriska studien. Beroende av deras relevansgrad hade de påverkat resultatet. Med tanke på detta ingår det en parallell uträkning i uppsatsen som visar hur resultatet hade blivit, å ena sidan, om alla dokument utan abstrakt varit relevanta, å andra sidan hur resultatet hade påverkats om alla dokument utan abstrakt varit icke-relevanta. Denna uträkning kallas för *Bästa/Sämsta tänkbara utfall* i den här uppsatsen.

Uträkningarna av P(10) och AP i Bästa/Sämsta tänkbara utfall, utfördes på samma sätt som det beskrivs ovan i avsnitt 4.3.2. Uträkningen utgick i första fallet, Bästa tänkbara utfall, från att de dokument som saknar abstrakt var relevanta och i andra fallet, Sämsta tänkbara utfall, från att de inte var relevanta. I den parallella uträkningen tog ett dokument utan abstrakt nu plats i träfflistan, varpå den påverkade placeringarna genom att skjuta på efterföljande dokument. När placeringarna ändrades påverkade detta resultatet av P(10) och AP. Dessutom påverkade förskjutningen i placeringarna den parallella uträkningen genom att dokument 20 blev dokument 21 och därmed inte togs med i beräkningen av AP.

Dokument på annat språk än engelska, som på grund av detta inte inkluderades i relevansbedömningen, ingick i Bästa tänkbara/Sämsta tänkbara utfall på samma sätt som dokument utan abstrakt.

Eftersom initialdokumentet inte skulle bidra med ny information för användaren så togs dock dessa dokument inte med i uträkningen av Bästa tänkbara/Sämsta tänkbara utfall. Samma sak gällde när dubletter återvanns, det vill säga om ett dokument dök upp en andra gång, eftersom användaren då redan fått tillgång till informationen i detta dokument.

## 5. Resultat

### 5.1 Återvinningsseffektivitet

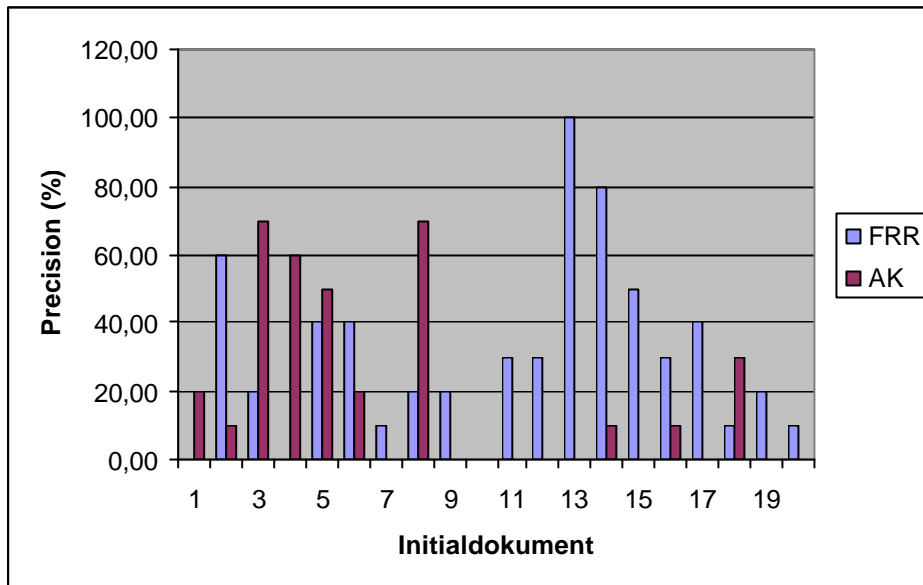
I det här kapitlet redovisas den empiriska studiens resultat i form av tabeller och diagram med kommentarer.

I de fall själva initialdokumentet återvanns bland dokumenten i träfflistan ersattes det av nästa dokument eftersom initialdokumentet inte bidrar med någon ny information. Detta skedde 6 gånger i AK. FRR-metoden kopplar ihop initialdokumentet med dokument som den delar en referens med, därför kan inte den här metoden återvinna initialdokumentet. Däremot är det möjligt att en dubblett, alltså samma dokument som initialdokumentet publicerat i en annan källa, återvinns. Detta skedde dock inte. Emellertid återvann FRR-metoden vid ett tillfälle 1 dubblett av ett dokument i samma träfflista. Dokumentet hade publicerats i två olika tidskrifter och fanns därmed som två olika dokument i samma träfflista.

Tabell 1 och figur 2 visar resultatet av P(10). Resultaten visar att FRR har ett bättre medelvärde vad gäller precision än AK med sina 30,50 % jämfört med 17,50 % för AK. Som tabellen visar så återvann AK-metoden inte ett enda relevant dokument bland de 10 första dokumenten i 10 fall av 20. Motsvarande siffra för FRR-metoden är 3 av 20. Det bästa resultatet som uppmättes, 100 % precision, hittar man vid initialdokument 13 i kolumnen för FRR. Där återvanns 10 relevanta dokument av 10 möjliga medan AK inte återvann några relevanta dokument alls vid samma initialdokument.

**Tabell 1.** P(10).

<b>Initialdokument</b>	<b>FRR (%)</b>	<b>AK (%)</b>
1	0,00	20,00
2	60,00	10,00
3	20,00	70,00
4	0,00	60,00
5	40,00	50,00
6	40,00	20,00
7	10,00	0,00
8	20,00	70,00
9	20,00	0,00
10	0,00	0,00
11	30,00	0,00
12	30,00	0,00
13	100,00	0,00
14	80,00	10,00
15	50,00	0,00
16	30,00	10,00
17	40,00	0,00
18	10,00	30,00
19	20,00	0,00
20	10,00	0,00
<b>Medelvärde (%)</b>	<b>30,50</b>	<b>17,50</b>



**Figur 2.** P(10).

Tabell 2 visar antalet relevanta dokument per initialdokument och metod. Den visar även poolen, alltså det gemensamma antalet kända och unika relevanta dokument. I några fall, till exempel vid initialdokument 5, återvann båda metoderna ett och samma relevanta dokument. Detta dokument räknas bara en gång, vilket förklarar att det i poolen finns 19 dokument i stället för 20 vid initialdokument 5.

**Tabell 2.** Antalet relevanta dokument fördelade efter metod.

Initial-dokument	Antal relevanta dokument genom FRR	Antal relevanta dokument genom AK	Totala antalet unika relevanta dokument i poolen
1	0	5	5
2	7	1	7
3	4	11	15
4	0	15	15
5	6	14	19
6	8	6	14
7	4	0	4
8	2	15	17
9	3	0	3
10	0	0	0
11	7	0	7
12	3	0	3
13	18	0	18
14	13	2	14
15	12	0	12
16	4	1	4
17	4	0	4
18	1	3	4
19	3	0	3
20	2	0	2

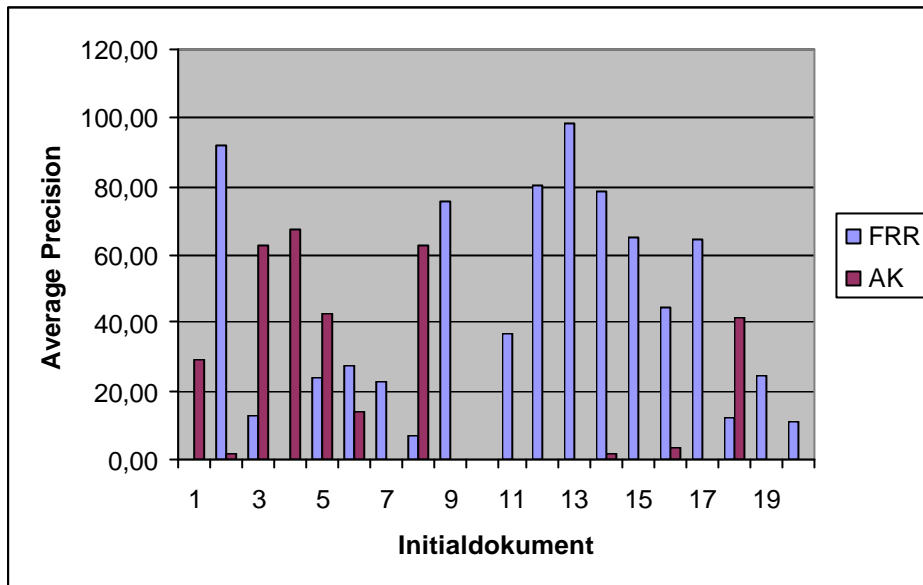
Tabell 3 och figur 3 visar AP. Precisionen har beräknats vid varje relevant dokument i listan av 20 återvunna dokument. Sedan har värdena summerats och summan har dividerats med antalet kända och unika relevanta dokument.

Resultatet visar att FRR presterar bättre än AK, med 38,93 % mot 16,39 % i medelvärde, MAP. Även här står FRR-metoden för det maximala värdet, 98,35 %, vid initialdokument 13, medan AK-metoden inte återvann ett enda relevant dokument vid samma initialdokument.

Vid beräkningarna av AP-värdet gällde  $DCV = 20$ . Trots detta gäller samma siffror vad gäller antalet sökningar som genererade 0 relevanta dokument för både P(10) och AP. FRR-metoden återvann inga relevanta dokument vid 3 av 20 gånger och AK-metoden återvann inte en enda relevant dokument i 10 fall av 20.

**Tabell 3.** AP.

<b>Initialdokument</b>	<b>FRR (%)</b>	<b>AK (%)</b>
1	0,00	29,41
2	91,96	1,59
3	12,90	62,84
4	0,00	67,61
5	23,60	42,68
6	27,49	13,96
7	22,44	0,00
8	7,06	62,63
9	75,76	0,00
10	0,00	0,00
11	36,89	0,00
12	80,56	0,00
13	98,35	0,00
14	78,46	2,01
15	65,33	0,00
16	45,00	3,57
17	64,59	0,00
18	12,50	41,52
19	24,44	0,00
20	11,25	0,00
<b>MAP (%)</b>	<b>38,93</b>	<b>16,39</b>



Figur 3. AP.

## 5.2 Signifikanstestning

För att se om det föreligger en signifikant skillnad mellan AP-resultaten för de båda metoderna FRR och AK har den här empiriska studiens resultat signifikanstestats. Genom detta ställs en nollhypotes,  $H_0$ , mot ett alternativ till denna, en så kallad mothypotes,  $H_1$ . Signifikanstestningen ska leda till att man antingen förkastar nollhypotesen eller accepterar den.<sup>90</sup> Signifikanstestningen i den här uppsatsen har utförts med hjälp av Wilcoxons teckenrangtest. Det är ett icke-parametriskt test som används för att pröva nollhypotesen vid parvisa observationer och baseras på rangtal.<sup>91</sup> Icke-parametriska test kan användas när involverade populationer inte kan antas vara normalfördelade och när stickprovet är litet, som fallet är i den här studien.<sup>92</sup> Wilcoxons teckenrangtest baseras på differensen mellan observationerna för varje par.<sup>93</sup>

Wilcoxons teckenrangtest går i praktiken ut på följande sätt. Man har  $n$  stycken observationspar där man beräknar differensen mellan mätvärdena. Differenserna mellan mätvärdena ordnas stigande efter deras absolutbelopp, alltså utan hänsyn till tecken. De par vars differens är lika med 0 stryks. Differenserna får olika rangtal, där minsta differensen får rangtalet 1. Rangsumman för de positiva respektive negativa differenserna räknas ihop. Den minsta av dessa två summor blir testfunktion i Wilcoxons test och skrivs  $T$ .<sup>94</sup>

När  $T$  är bestämt jämförs det med det kritiska värdet på  $T$  för Wilcoxons teckenrangtest. Kritiska värdet återfinns i en tabell och fås fram genom att titta på signifikansnivån, som bestämts i förväg, och antalet observationspar,  $n$  stycken. Signifikansnivån,  $\alpha$ , för detta

<sup>90</sup> Körner, Svante & Wahlgren, Lars (2000). *Statistisk dataanalys*, s. 173.

<sup>91</sup> Gravetter, Frederick J. & Wallnau, Larry B. (2000). *Statistics for the Behavioral Sciences*, s. 649.

<sup>92</sup> Körner & Wahlgren (2000), s. 304.

<sup>93</sup> Gravetter & Wallnau (2000), s. 649.

<sup>94</sup> Gravetter & Wallnau (2000), s. 649-651.

signifikanstest bestämdes till 5 %. Nollhypotesen förkastas om  $T$  är mindre än eller lika med det kritiska värdet, i annat fall behålls den.<sup>95</sup>

Signifikanstestningen utgick från följande hypoteser:

$H_0$ : de båda metoderna är lika effektiva med avseende på AP.

$H_1$ : de båda metoderna är inte lika effektiva med avseende på AP.

Tabell 4 visar hur resultatet av de båda sökmetodernas AP-värden signifikanstestades med hjälp av Wilcoxons teckenrangtest. I tabellen har AP-värdenas differens räknats ut. Sedan har differensen omvandlats till absoluta differensvärden, det vill säga, värdena skrivs ut oberoende av tecken. Dock har de positiva värdena markerats i fet stil i tabellen för att skilja dem från de negativa värdena och på så sätt underlätta för läsaren. Slutligen har värdena rankats från det minsta värdet till det största. Initialdokument 10 stryks på grund av att differensen är 0, vilket förklarar att rankningen går från 1 till 19 trots att det är 20 initialdokument.

**Tabell 4.** Wilcoxons teckenrangtest.

Initialdokument	FRR (%)	AK (%)	Differens	Abs.differens	Rankning
10	0,00	0,00	0,00	0,00	0
20	11,25	0,00	-11,25	11,25	1
6	27,49	13,96	-13,53	13,53	2
5	23,60	42,68	19,08	<b>19,08</b>	<b>3</b>
7	22,44	0,00	-22,44	22,44	4
19	24,44	0,00	-24,44	24,44	5
18	12,50	41,52	29,02	<b>29,02</b>	<b>6</b>
1	0,00	29,41	29,41	<b>29,41</b>	<b>7</b>
11	36,89	0,00	-36,89	36,89	8
16	45,00	3,57	-41,43	41,43	9
3	12,90	62,84	49,94	<b>49,94</b>	<b>10</b>
8	7,06	62,63	55,57	<b>55,57</b>	<b>11</b>
17	64,59	0,00	-64,59	64,59	12
15	65,33	0,00	-65,33	65,33	13
4	0,00	67,61	67,61	<b>67,61</b>	<b>14</b>
9	75,76	0,00	-75,76	75,76	15
14	78,46	2,01	-76,45	76,45	16
12	80,56	0,00	-80,56	80,56	17
2	91,96	1,59	-90,37	90,37	18
13	98,35	0,00	-98,35	98,35	19

Rankningarna för de positiva värdena, 3, 6, 7, 10, 11, och 14, har summerats. Resultatet blir  $R=51$ . Samma sak utförs med de negativa värdenas rankningar, vilket ger  $R=139$ .

Den minsta av dessa summer är 51 och detta värde blir alltså värdet på testfunktionen  $T$  i Wilcoxons teckenrangtest.

I en tabell med kritiska värden av  $T$  för Wilcoxons teckenrangtest, letas den siffra upp som motsvarar signifikansnivån 5 % i relation till antalet parvisa observationer,  $n$ , i det

<sup>95</sup> Gravetter & Wallnau (2000), s. 650-651.

här fallet 19.<sup>96</sup> Det kritiska värdet i det här fallet blir 46.  $H_0$  förkastas om  $T$  är mindre än eller lika med det kritiska värdet. Eftersom  $T = 51$ , ett större värde än 46, accepteras  $H_0$ .

Genom att acceptera  $H_0$  förkastas  $H_1$ . Resultatet av den empiriska studien *kan* alltså bero på slumpen, resultatet är inte signifikant, det är inte statistiskt säkerställt.

### 5.3 Bästa/Sämsta tänkbara utfall

När träfflistorna granskades var det sammanlagt 11 dokument vars abstrakt inte gick att finna, 10 i FRR och 1 i AK. Dessutom återvanns ett dokument på annat språk än engelska i FRR. Dessa träffar gick följaktligen inte att relevansbedöma utan ersattes av nästa träff i listan. På grund av detta har en parallell uträkning genomförts där de träffar som saknade abstrakt/var på ett annat språk än engelska dels antagits vara relevanta, dels icke-relevanta.

Följande tabell, tabell 5, visar resultatet av uträkningen av  $P(10)$  vid Bästa tänkbara utfall. Som man kan se skiljer sig inte resultaten nämnvärt från värdena i den ursprungliga uträkningen i 5.1. Medelvärdet för  $P(10)$  i den ursprungliga uträkningen var 30,50 % i FRR och 17,50 % i AK, vilket kan jämföras med medelvärdet för Bästa tänkbara utfall, 31,50 % och 18,00 % i FRR respektive AK. De värden som påverkades är markerade i fet stil.

**Tabell 5.**  $P(10)$  vid Bästa tänkbara utfall.

Initialdokument	FRR (%)	AK (%)
1	<b>10,00</b>	20,00
2	60,00	10,00
3	20,00	70,00
4	<b>10,00</b>	60,00
5	40,00	50,00
6	40,00	20,00
7	10,00	0,00
8	20,00	70,00
9	20,00	0,00
10	0,00	0,00
11	30,00	0,00
12	30,00	<b>10,00</b>
13	100,00	0,00
14	80,00	10,00
15	50,00	0,00
16	30,00	10,00
17	40,00	0,00
18	10,00	30,00
19	20,00	0,00
20	10,00	0,00
<b>Medelvärde (%)</b>	<b>31,50</b>	<b>18,00</b>

<sup>96</sup> I denna studie har man utgått från tabell B.10 i Gravetter & Wallnau (2000), s. 704.

Tabell 6 presenterar värdena av P(10) i ett Sämsta tänkbara utfall. Som synes är det bara ett värde, markerat i fet stil, som har ändrats om man jämför med den ursprungliga uträkningen. Följaktligen har medelvärdet inte ändrats särskilt mycket. I den ursprungliga uträkningen var medelvärdet 30,50 % i FRR och 17,50 % i AK att jämföras med 30,00 % och 17,50 % i FRR respektive AK i Sämsta tänkbara utfall.

**Tabell 6.** P(10) vid Sämsta tänkbara utfall.

<b>Initialdokument</b>	<b>FRR (%)</b>	<b>AK (%)</b>
1	0,00	20,00
2	60,00	10,00
3	20,00	70,00
4	0,00	60,00
5	40,00	50,00
6	40,00	20,00
7	10,00	0,00
8	20,00	70,00
9	20,00	0,00
10	0,00	0,00
11	<b>20,00</b>	0,00
12	30,00	0,00
13	100,00	0,00
14	80,00	10,00
15	50,00	0,00
16	30,00	10,00
17	40,00	0,00
18	10,00	30,00
19	20,00	0,00
20	10,00	0,00
<b>Medelvärde (%)</b>	<b>30,00</b>	17,50



Tabell 7 redovisar resultatet av AP vid Bästa tänkbara utfall. MAP-värdet i den ursprungliga uträkningen var 38,93 % för FRR och 16,39 % för AK. I Bästa tänkbara utfall däremot var siffrorna 37,96 % för FRR och 15,66 % för AK. MAP-värdet här skiljer sig alltså inte markant från MAP-värdet i den ursprungliga uträkningen. Däremot kan tydliga avvikelser noteras vid enskilda AP-värden. Dessa värden är markerade i fet stil.

**Tabell 7.** AP vid Bästa tänkbara utfall.

<b>Initialdokument</b>	<b>FRR (%)</b>	<b>AK (%)</b>
1	<b>6,01</b>	<b>18,38</b>
2	<b>86,02</b>	<b>1,39</b>
3	12,90	62,84
4	<b>0,69</b>	<b>63,39</b>
5	<b>24,60</b>	<b>40,55</b>
6	27,49	13,96
7	<b>23,83</b>	0,00
8	7,06	62,63
9	75,76	0,00
10	0,00	0,00
11	<b>39,71</b>	0,00
12	<b>55,00</b>	<b>2,86</b>
13	<b>98,53</b>	0,00
14	78,46	2,01
15	<b>65,31</b>	0,00
16	45,00	3,57
17	64,59	0,00
18	12,50	41,52
19	24,44	0,00
20	11,25	0,00
<b>MAP (%)</b>	<b>37,96</b>	<b>15,66</b>

Tabell 8 redogör för AP vid Sämsta tänkbara utfall. I den ursprungliga uträkningen var MAP för FRR 38,93 % och 16,39 % för AK. Vid Sämsta tänkbara utfall är siffrorna 38,74 % för FRR och 16,39 % för AK. Här är skillnaderna inte heller så stora jämfört med den ursprungliga uträkningen. Endast vid fyra AP-värden, markerade i fet stil, går det att se en förändring.

**Tabell 8.** AP vid Sämsta tänkbara utfall.

<b>Initialdokument</b>	<b>FRR (%)</b>	<b>AK (%)</b>
1	0,00	29,41
2	91,96	1,59
3	12,90	62,84
4	0,00	<b>67,51</b>
5	<b>23,47</b>	42,68
6	27,49	13,96
7	22,44	0,00
8	7,06	62,63
9	75,76	0,00
10	0,00	0,00
11	<b>34,67</b>	0,00
12	80,56	0,00
13	<b>96,89</b>	0,00
14	78,46	2,01
15	65,33	0,00
16	45,00	3,57
17	64,59	0,00
18	12,50	41,52
19	24,44	0,00
20	11,25	0,00
<b>MAP (%)</b>	<b>38,74</b>	16,39

## 6. Diskussion

### 6.1 Relevansbedömning

Relevansbedömningen av träfflistorna gjordes enskilt av uppsatsens båda författare. Tanken bakom detta var att få en så objektiv bedömning som möjligt. Relevansgranskningen började med att initialdokumentets titel och abstrakt lästes igenom för att skapa en bild av vad dokumentet handlar om. Utifrån detta formades sedan ett informationsbehov, alltså ett behov av liknande dokument inom exakt samma ämnesområde som initialdokumentet. För varje initialdokuments informationsbehov fastställdes ett relevanskriterium som skulle fungera som mall vid relevansgranskningen av de nya dokumenten som de båda sökmetoderna genererade. Relevanskriterierna diskuterades av författarna för att deras innebörd skulle vara tydlig för båda. Efter en enskild granskning av träfflistorna jämfördes resultatet och eventuella diskrepanser diskuterades för att få en slutgiltig och överensstämmande bedömning.

Ett protokoll fördes över de avvikelser som förekom i relevansbedömningen och resultatet av detta blev en överensstämmelse på 96,25 %. Bland de 800 dokument som relevansgranskades förekom det skillnader i bedömningen i 30 fall. Trots en noggrann genomgång av relevanskriterierna förekom det alltså avvikelser i bedömningen och man kan fråga sig vad detta beror på. Avvikelserna kan dels bero på den mänskliga faktorn, det ligger alltid subjektivitet bakom bedömningar, dels på att trots att kriterierna noggrant granskats och diskuterats så kan missuppfattningar ändå ske. Detta kan bero på hur kriteriet formulerats och hur bedömarna tolkar vissa ord och uttryck.

Vid relevansbedömningen användes en binär relevansskala, en skala som kan kritiseras för att vara alldeles för strikt eftersom dokument anses vara antingen relevanta eller icke-relevanta. Olika grader av relevans tas alltså ej i beaktande. När det gäller relevansbedömningen i den här uppsatsen så är en binär skala dock bra, med tanke på att en flergradig skala hade ökat risken för större subjektivitet i bedömningen och haft som följd att diskrepansen mest troligt blivit större.

Subjektivitet och avvikelser i en bedömning kan kanske ses som ett problem men så länge ingen metod favoriseras blir subjektiviteten inget större bekymmer. Vid fastställandet av kriterierna klipptes initialdokumentens titel och abstrakt in i ett särskilt dokument. Author Keywords-fältet fanns alltså inte tillgå, detta för att inte gynna AK-metoden. I klassisk pearl growing är det dock en bra idé att titta på de nyckelord som tilldelats ett dokument och använda dessa som söktermer om man är intresserad av att finna dokument inom samma ämnesområde som initialdokumentet.

Man kan fråga sig hur den empiriska studiens resultat hade sett ut om någon annan person konstruerat relevanskriterierna och/eller utfört bedömningarna, med tanke på subjektiviteten. Det finns en möjlighet att resultatet hade sett annorlunda ut men den här studien utgår från fastställda kriterier som inte gynnar någon metod och därför ses inte detta som något problem.

## 6.2 Resultatet

De mått som användes i uppsatsen för att mäta effektiviteten av sökmetoderna FRR och AK var P(10) och AP. Måtten kompletterar varandra då P(10) beräknar precision, något som AP också gör men med en inbyggd recall-aspekt. AP-metoden utnyttjar en pool med de relevanta dokumenten från båda sökmetoderna, dokument som antas vara relevanta för sökningen.

### 6.2.1 Precision (10)

Beräkningarna av P(10) resulterade i att FRR presterade 30,50 % och AK presterade 17,50 % i medelvärde. Det skiljer sig ganska mycket i medelvärde mellan metoderna, där FRR presterar bättre. I 13 fall av 20 presterade FRR-metoden bättre än AK-metoden som i sin tur uppvisade ett bättre resultat än FRR-metoden i 6 fall av 20. Vid ett fall, initialdokument 10, återvann ingen av metoderna några relevanta dokument alls.

AK-metoden återvann inte ett enda relevant dokument bland de tio första dokumenten hälften av gångerna, det vill säga vid 10 fall av 20. Detta kan härledas till det faktum att författarnas nyckelord för det mesta är väldigt breda och då återvinner information från andra, icke-relevanta, ämnesområden. Ett exempel som genom AK-metoden genererade 0 relevanta dokument var initialdokument 19 som handlar om bias när det gäller information på Internet. Relevanskriteriet för sökningen var dokument som problematiserar Internet som informationskälla. Författarens två första nyckelord var *information* och *Internet*, termer som innefattar väldigt mycket och som inte nödvändigtvis behöver vara kopplade till det ämnesområde som den här uppsatsen fokuserar på, det vill säga B&I. AK-metoden, såsom den används i den empiriska studien, utgår från de två första nyckelorden i AK-fältet. Det innebär att eventuellt bättre sökresultat hade kunnat fås om bättre nyckelord, placerade längre bak i Author Keywords-fältet, hade utnyttjats.

Genom FRR-metoden söks det inte med hjälp av söktermer, i stället kopplas dokument ihop genom sina referenslistor. Den här metoden undviker alltså de svårigheter som alltför breda/små söktermer bär med sig. Detta kan man se i resultatlistan, där FRR endast i 3 fall av 20 inte återvann några relevanta dokument alls bland de 10 första dokumenten. Ett av dessa fall var det tidigare nämnda initialdokument 10, som uppvisade 0 % i precision vid sökningar genom båda metoderna. Man ska dock komma ihåg att 0 också är ett resultat. Vad som kan göras är att undersöka vad resultatet beror på.

Skälet till att FRR presterade dåligt i initialdokument 10 kan bero på att kopplingsstyrkan inte var så stark, den delade endast en referens med de dokument den hade bibliografisk koppling till. Ju fler gemensamma referenser som delas mellan dokument desto större är chansen att dokumenten har någonting gemensamt, en teori som bland annat forskarna Sen och Gan arbetat med.<sup>97</sup> I detta fall kan man tydligt se att teorin stämmer.

---

<sup>97</sup> Sen & Gan (1983), s. 78.

När det kommer till AK-metoden för initialdokument 10 så var de två första nyckelorden, *information needs* och *information resources*, väldigt breda. Eftersom initialdokumentet behandlade skolbibliotek och dess utveckling i Turkiet fokuserades relevanskriteriet kring just utvecklingen av skolbibliotek. Något specifikt land kopplades inte till relevanskriteriet för att inte begränsa detta för mycket. Trots det tillåtande relevanskriteriet återvanns inga nya dokument inom ämnet och, som redan nämnts, beror detta på att de två första nyckelorden som givits initialdokumentet är breda. Nyckelord nummer tre, *school libraries*, hade kanske bidragit till att generera fler relevanta dokument.

Ett annat intressant fall är initialdokument 13 där resultatet i P(10) blev 100 % för FRR och 0 % för AK. Resultatet av FRR i detta fall var det bästa resultatet som uppmättes i den empiriska studien. Initialdokumentet handlade om Impact Factor och därmed var relevanskriteriet fokuserat kring detta. Att FRR ger ett så bra resultat kan bero på att kopplingsstyrkan är relativt stark, det första återvunna dokumentet delar 9 referenser med initialdokumentet och nästa dokument delar 6 referenser med initialdokumentet. Av större vikt kan dock vara att dessa dokument är skrivna av samma författare som initialdokumentets och kan därmed ge en förklaring till varför de har så många gemensamma referenser. Det sista dokumentet som relevansgranskades i träfflistan för initialdokument 13 hade en kopplingsstyrka på 3, initialdokumentet har med andra ord minst 3 gemensamma referenser med de dokument den är bibliografiskt kopplad till.

Skälet till att AK-metoden vid initialdokument 13 inte genererade några relevanta dokument alls bland de 10 första kan bero på att författarens nyckelord var *bibliometrics* och *citation analysis*. Eftersom de dokument som efterfrågades skulle handla om Impact Factor så blev söktermerna alltför breda och genererade inga relevanta dokument. Först på tredje plats hade författaren skrivit *impact factor* som nyckelord och därmed var den termen inte till någon hjälp i den här sökningen.

Generellt sett så har uträkningen av P(10) inte gett något tillfredsställande resultat för någon av sökmetoderna. Båda metoderna har endast i 4 fall av 20 genererat 5 eller fler relevanta dokument bland de 10 första, alltså en precision på 50 % eller mer. En hög precision är önskvärd eftersom användare ofta nöjer sig med att granska träfflistans första sida som ofta bara innehåller 10 stycken dokument.

### 6.2.2 Uninterpolated Average Precision

Resultatet visar att FRR-metoden presterar bättre än AK-metoden, med 38,93 % mot 16,39 % i medelvärden, MAP. Även i uträkningen av AP står FRR-metoden för de maximalt enskilda värdena, 98,35 % och 91,96 %, vid initialdokument 13 och 2, medan AK-metoden visar 0 % och 1,59 % vid samma initialdokument.

Initialdokument 2 handlar om optimering av booleska sökfrågor med hjälp av genetiska algoritmer. De två första söktermerna i AK är *evolutionary algorithms* och *genetic algorithms*. Med tanke på att algoritmer av detta slag används inom olika områden, inte bara inom B&I, så visade träfflistan dokument som behandlade ämnen inom vitt skilda områden. Följden av detta blev att AK-metodens resultat i detta fall var lågt, 1,59 %.

FRR-metoden genererade, som redan nämnts, 91,96 % i resultat för initialdokument 2. Det här fallet visar på ett exempel där FRR-metoden fungerar på ett bra sätt. Efter en granskning av träfflistan på dokumenten som delar referenser med initialdokument 2 kan man se att de 6 första dokumenten i träfflistan är relevanta. Det förekommer även ett relevant dokument till som finns placerat på plats 16 i träfflistan. De höga placeringarna av de relevanta träffarna förklarar resultatet. De 3 första dokumenten i träfflistan har dessutom samma författare, dock inte samma författare som initialdokumentet, och har en kopplingsstyrka på 4, de har alltså 4 gemensamma referenser med initialdokumentet. De relevanta dokumenten har till stor del använt sig av samma referenser, även om det förekommer individuella skillnader. Den här granskningen visar på att författare ofta använder sig av vissa centrala verk inom sitt område och då fungerar bibliografisk koppling bra. Eftersom algoritmer, som ingår i informationsbehovet, dock används inom många områden hamnar en hel del icke-relevanta dokument i listan. Dessa dokument delar nästan uteslutande samma två verk som handlar om algoritmer men som inte tvunget kopplas till B&I.

AP-uträkningen gav inte 100 % i resultat vid något enskilt fall. Initialdokument 13 fick 100 % i FRR i P(10) eftersom alla de tio första dokumenten var relevanta. I AP-uträkningen blev resultatet istället 98,35 %. När det kommer till AP måste man ta hänsyn till poolen vid uträkningen. Poolen består av alla relevanta dokument, när det gäller initialdokument 13 är poolens storlek 18. AK-metoden genererade inte några relevanta dokument alls varpå den inte påverkade poolen i det här fallet. Eftersom placeringarna av de relevanta dokumenten spelar in i uträkningen av AP så uppnås inte 100 % här eftersom ett icke-relevant dokument finns bland de 18 första placeringarna. 100 % kan uppnås i AP om så endast ett enda relevant dokument återvinns av någon av metoderna. För att 100 % ska uppnås ska det relevanta dokumentet dock finnas på första plats i träfflistan och den andra sökmetoden ska inte återvinna ett enda relevant dokument, den ska alltså inte påverka poolen som består av ett dokument.

Initialdokument 4 visade AK-metodens bästa resultat med 67,61 % mot 0 % i FRR-metoden. Initialdokumentet handlade om klassifikation i samband med informationsåtervinning vilket också blev relevanskriteriet vid bedömningen av träfflistorna. De nyckelord som användes i AK-metoden var *information retrieval* och *classification*. Söktermerna var mycket lämpliga med tanke på initialdokumentets innehåll, vilket också visade sig i resultatet.

FRR-metodens dåliga resultat vid initialdokument 4 kan förklaras efter en granskning av FRR-metodens träfflista. Där kan man se att kopplingsstyrkan är 3 för de 3 första dokumenten och 2 för de resterande 17 dokumenten, kopplingsstyrkan är med andra ord relativt låg. Vad som också ska uppmärksammas är att alla 20 dokument delar två specifika referenser med initialdokumentet, referenser med samma författare. Dessa referenser behandlar dock ett område som inte har någon direkt koppling till B&I utan till ämnet psykometri. Referenserna har bevisligen en koppling till initialdokumentets ämne eftersom de finns med i dess referenslista men inte så starkt att den bibliografiska kopplingen återvinner dokument inom exakt samma ämnesområde.

Avslutningsvis kan man konstatera att samma siffror gäller när det kommer till antalet sökningar som genererade 0 relevanta dokument för både P(10) och AP. FRR-metoden återvann inga relevanta dokument 3 av 20 gånger och AK-metoden återvann inte ett enda relevant dokument i 10 fall av 20. Med tanke på att AP räknas ut vid  $DCV = 20$

hade det kunnat vara så att uträkningen av AP gett ett bättre resultat än P(10) om sökmetoderna återvunnit relevanta dokument i positionerna 11-20 även om de inte återvunnit några i positionerna 1-10. Nu hände inte detta i den här studien, vilket förklarar varför det är lika många 0 % -värden i både P(10) och AP.

### 6.2.3 Pearl growing

De två metoderna som testas och jämförs i den här uppsatsen är båda två exempel på pearl growing-metoder. Båda metoderna utgår från ett initialdokument och försöker generera nya dokument inom samma ämnesfält. Dock finns det skillnader metoderna emellan.

När AK-metoden, såsom den tillämpas i den här uppsatsen, används som pearl growing-metod är den beroende av vilka nyckelord som författaren valt ut till initialdokumentet och i vilken ordning de står, särskilt om metoden skulle implementeras som automatisk tjänst. I den här uppsatsen har enbart de två första nyckelorden använts som söktermer och därför har ordningen många gånger visat sig ha en avgörande betydelse. Vid en automatisering skulle dokumentens författare behöva lägga ner mer tid på att verkligen välja ut lämpliga nyckelord och dessutom placera dem i prioriteringsordning för att tjänsten skulle fungera på bästa möjliga sätt. Allt detta för att underlätta sökningen efter likartade dokument. AK-metoden som pearl growing-metod är begränsad till sin utformning. Inom klassisk pearl growing har man tillgång till hela Author Keywords-fältet samt bland annat även titel och abstrakt. Detta underlättar pearl growing-processen avsevärt.

I resonemanget kring nyckelorden kan man även fundera över hur resultatet sett ut om studien utförts inom ett naturvetenskapligt område. Termer inom naturvetenskap är ofta mer pregnanta än de termer som används inom samhällsvetenskap och humaniora. Detta hade kanske kunnat leda till mer exakta nyckelord att söka på och därmed ett bättre resultat.

Ett annat fält som i ISI Web of Science skulle kunna användas i en klassisk pearl growing är KeyWords Plus. Detta fält i den fullständiga bibliografiska posten innehåller ord och fraser som frekvent förekommer i titlarna på dokumentets referenser men som inte nödvändigtvis förekommer i dokumentets titel och/eller Author Keywords-fältet. Ett av syftena med KeyWords Plus-fältet kan vara att ge förslag till nya söktermer för att finna fler likartade dokument. Detta fält kan dock vara ett problem när man söker på nyckelord bland Author Keywords, såsom det gjorts i den här uppsatsen, eftersom systemet återvinner dokument där söktermerna återfinns någonstans i den bibliografiska posten. I de fall då termerna finns i KeyWords Plus kan det bidra till att icke-relevanta dokument återvinns eftersom KeyWords Plus inte hänvisar till själva dokumentet utan till dess referenser.

Till skillnad från AK-metoden kopplar FRR-metoden ihop initialdokumentet med dokument som det har minst en gemensam referens med. I den här uppsatsen har FRR-metoden uppvisat ett bättre resultat än AK-metoden, varpå man kan säga att FRR-metoden här presterat bättre som pearl growing-metod. FRR-metoden har dock sina tillkortakommanden. En nackdel kan vara att när det kommer till de gemensamma referenserna så får man betänka att dessa till exempel kan bestå av dokument som

behandlar teorier och/eller metoder, eller andra dokument som inte direkt har något med dokumentets ämne att göra. Dokumenten som har samma referens kan alltså behandla vitt skilda ämnen varpå de nya dokumenten inte kan anses vara relevanta.

Kopplingsstyrkan brukar anses vara av stor vikt när det kommer till bibliografisk koppling.<sup>98</sup> I vissa fall visar en stark kopplingsstyrka på en stor samhörighet mellan de dokument som har gemensamma referenser. I många fall dock, enligt den här uppsatsens resultat, har inte kopplingsstyrkan någon betydelse alls. Kopplingsstyrkan är alltså inte alltid att lita på och därmed kan den inte alltid användas som ett mått på relevans. Kopplas inte initialdokumentet till relevanta dokument så fungerar inte FRR-metoden i pearl growing-sammanhang.

### 6.3 Utfallet av signifikanstestningen

Med tanke på att FRR-metoden presterar mycket bättre än AK-metoden både när det gäller P(10) och AP ansågs det intressant att se om det gick att statistiskt säkerställa resultatet. Ett signifikanstest utfördes på AP-värdena med hjälp av Wilcoxons teckenrangtest.

Signifikanstestets resultat visade att det inte fanns någon signifikans.  $T = 51$  är större än det kritiska värdet för signifikansnivån 5 % som är 46. Eftersom det inte förekom någon signifikans behålls  $H_0$ , det vill säga hypotesen som säger att metoderna är lika effektiva med avseende på AP. Samtidigt förkastas  $H_1$  och det är därför omöjligt att dra den generella slutsatsen att metoderna inte är lika effektiva med avseende på precision.

Ett av skälen till att ingen signifikans erhöles kan vara att resultaten i precisionsberäkningarna inte är konsekventa. Medelvärdet för FRR är klart bättre än medelvärdet för AK både vid P(10) och vid AP. Vid en granskning av hur sökmetoderna presterar vid varje enskild sökning ser man att AK-metoden presterar bättre än FRR vid 6 tillfällen samtidigt som den får 0 % i precision vid 10 tillfällen. FRR å sin sida presterar bättre än AK vid 13 tillfällen samtidigt som den får 0 % i precision vid 3 tillfällen. För att få signifikans hade FRR-metoden varit tvungen att prestera bättre än AK-metoden i ytterligare ett antal fall.

### 6.4 Bästa/Sämsta tänkbara utfall

Eftersom några av de dokument som återvanns i den empiriska studien inte inkluderade ett abstrakt eller var på annat språk än engelska var dessa inte möjliga att relevansgranska. De ersattes då av nästkommande dokument med abstrakt i listan. En parallell uträkning utfördes därför i uppsatsen för att se om och hur resultatet hade påverkats om abstrakt funnits eller om alla dokumentet varit på engelska. I Bästa tänkbara utfall antogs alla dokumenten utan abstrakt vara relevanta och i Sämsta tänkbara utfall icke-relevanta.

---

<sup>98</sup> Sen & Gan (1983), s. 78.



AK-metoden avgränsades till engelska, detta för att undvika dokument på andra språk än engelska och som därmed varit omöjliga att relevansgranska. Hade inte denna begränsning gjorts kunde resultatet ha blivit lidande av att för många dokument hade fått ersättas. För FRR-metoden var det inte möjligt att avgränsa språket men turligt nog återvanns endast ett dokument på annat språk än engelska, vilket fick ersättas med nästa dokument i listan. Dokumentet på annat språk än engelska togs sedan med i uträkningen av Bästa/Sämsta tänkbara utfall. Som nämnts ovan så visar det sig att resultatet inte blev speciellt annorlunda även om de dokument som ersatts tagits med i den ursprungliga uträkningen.

#### **6.4.1 Precision (10)**

Medelvärdet för P(10) i den ursprungliga uträkningen var 30,50 % för FRR-metoden och 17,50 % för AK-metoden. Jämför man siffrorna med medelvärdena för P(10) i Bästa tänkbara utfall så visar sig FRR få 31,50 % och AK får 18 %. Båda metoderna höjs med 1 respektive 0,5 procentenheter men det blir inga radikala skillnader. Alltså favoriserades ingen metod av att vissa dokument har fått ersättas med efterföljande dokument.

I Sämsta tänkbara fall blev medelvärdet för P(10) 30,00 % för FRR och 17,50 % för AK. Jämförs dessa resultat med den ursprungliga uträkningen ser man att FRR sänks med 0,5 procentenheter medan AK-metodens resultat inte förändras alls. Inte heller i Sämsta tänkbara utfall missgynnas någon av metoderna av att vissa dokument har fått ersättas med efterföljande dokument.

FRR-resultatet förändras till viss del i uträkningarna av P(10) i Bästa/Sämsta tänkbara utfall. Detta beror på att 4 av de 10 dokument som återvanns utan abstrakt av FRR befann sig i positionerna 1-10 och fick ersättas av nästföljande dokument.

Skälet till att AK-metodens resultat i P(10) inte förändrades i Sämsta tänkbara utfall beror på att endast ett dokument bland alla återvunna dokument genom AK-metoden saknade abstrakt, detta skedde vid initialdokument 12. Med tanke på att den sökningen inte återvann några relevanta dokument över huvud taget i träfflistan så förändrades inte resultatet om dokumentet utan abstrakt betraktades som icke-relevant. I Bästa tänkbara utfall förändrades däremot resultatet eftersom dokumentet där betraktades som relevant och gav den sökningen en ökning från 0 % till 10 %.

#### **6.4.2 Uninterpolated Average Precision**

MAP, medelvärdet för AP, var i den ursprungliga uträkningen 38,93 % för FRR och 16,39 % för AK. Jämför man dessa värden med siffrorna i Bästa tänkbara utfall så blir de 37,96 % för FRR och 15,66 % för AK. I Sämsta tänkbara utfall blir MAP för FRR 38,74 % och 16,39 % för AK. Medelvärdet har alltså inte förändrats nämnvärt. Dock är det intressant att titta på vissa enskilda AP-värden där större förändringar har ägt rum.

Vid FRR i initialdokument 1, till exempel, tillkommer det tre stycken relevanta dokument vid Bästa tänkbara utfall, något som inverkar på resultatet, som blir 6,01 %. I den ursprungliga uträkningen räknar man inte in några relevanta dokument i FRR och

därför blir resultatet 0 % där. Vad gäller AK i initialdokument 1 så påverkas det i motsatt riktning eftersom det inte förekommer några dokument utan abstrakt. Poolen av relevanta dokument ökar däremot med tre, från fem till åtta. Eftersom man då tvingas dividera med ett större tal så minskar AP-värdet från 29,41 % till 18,38 %.

Ytterligare ett exempel som är värt att ta upp är initialdokument 12. Där minskar AP-värdet i FRR från 80,56 % i den ursprungliga uträkningen till 55 % i Bästa tänkbara utfall. Värdet vid AK ökar dock från 0 % till 2,86 %. Sökningarna vid initialdokument 12 är de enda där dokument utan abstrakt återvunnits genom båda metoderna, en per metod. Poolen ökar med två dokument, vilket är orsaken till den stora förändringen.

De enskilda värdena för AP i Sämsta tänkbara utfall förändras inte lika mycket som vid Bästa tänkbara utfall. Vid Sämsta tänkbara utfall räknas dokument utan abstrakt som icke-relevanta. Ett icke-relevant dokument påverkar inte poolen så länge den inte påverkar placeringarna i träfflistan så att ett relevant dokument skjuts ut ur listan och poolen därmed blir mindre. Detta skedde dock inte i den här uppsatsen. I många fall där dokumenten saknade abstrakt fanns inga relevanta dokument i träfflistan från början, varpå ännu ett icke-relevant dokument inte hade någon betydelse. De små skillnaderna i resultat som finns beror dock på att placeringarna ändrades och fick konsekvenser för uträkningen.

Avslutningsvis kan man konstatera att trots att ett antal återvunna dokument inte tas med i den ursprungliga uträkningen så påverkar inte det här resultatet speciellt mycket, vare sig man antar att de dokument som inte tas med i den empiriska studien hade varit relevanta eller icke-relevanta.

## 6.5 Koppling till tidigare forskning

Den forskning inom bibliografisk koppling som presenterats i uppsatsen kommer här att diskuteras i relation till den här uppsatsens resultat.

**Kessler** introducerade begreppet bibliografisk koppling och skapade två kriterier för koppling. För den här uppsatsen är Kesslers kriterium A relevant.

Kriterium A: Ett antal dokument utgör en relaterad grupp  $G_A$  om varje dokument i gruppen har åtminstone en kopplingsenhet till ett givet testdokument  $P_o$ . Kopplingsstyrkan mellan  $P_o$  och något dokument ur  $G_A$  mäts genom antalet kopplingsenheter ( $n$ ) dem emellan.  $G_A^n$  är proportionen av  $G_A$  som är länkad till  $P_o$  genom  $n$  kopplingsenheter.<sup>99</sup>

$G_A$  kan sägas motsvara dokumenten i ISI Web of Science som har minst en koppling till  $P_o$ , initialdokumenten.

Kessler sammanfattar den bibliografiska kopplingens fem egenskaper som återvinningsverktyg på nedanstående sätt. Egenskaperna kommenteras här för att kopplas ihop med uppsatsen.

---

<sup>99</sup> Kessler (1962), s. 10.

- Bibliografisk koppling är oberoende av ord och språk. Därmed undviks språkliga svårigheter som syntax och språkvanor.

Man behöver inte behärska ämnesdisciplinens ordförråd eftersom bibliografisk koppling som sökmetod inte är beroende av söktermer. Man bör dock komma ihåg att koppling kan göras till dokument på alla möjliga språk, vilket kan bli ett problem vid relevansgranskningen om man inte behärskar språket som dokumentet är skrivet på.

- Det behövs ingen expertbedömning av dokumenten.

En viss förkunskap inom ämnesfältet behövs dock för att kunna genomföra en relevansbedömning av dokumenten och se om kopplingen har någon relevans till exempelvis ett visst informationsbehov.

- Bibliografisk koppling sträcker sig till det förflutna lika mycket som till framtiden.

Ett initialdokument  $P_0$  kan vara kopplat till ett dokument  $d_1$  som publicerats före  $P_0$  likaväl som till ett dokument  $d_2$  som publicerats efter  $P_0$ . Bibliografisk koppling är alltså inte tidsmässigt statiskt.

- Metoden producerar inte en statisk klassifikation för en given vetenskaplig artikel. Den genomgår förändringar som reflekterar den aktuella användningen av det vetenskapliga ämnesfältet.

Vad gäller den här fjärde egenskapen kan man ibland urskilja vilka källor som värderas högt och används flitigt inom ett visst område för tillfället. Bibliografisk koppling visar hur olika dokument värderats som informationskällor genom tiderna.

- Artiklar som delar en grupp av kopplingar med en given artikel kan betraktas som dess logiska referens och ses som ett substitut för dess egna referenser.  $G_A$  kan ses som  $P_0$ :s logiska referenser.<sup>100</sup>

Det är inte alltid detta stämmer om man utgår från informationsbehov. Det är också viktigt att tänka på syftet till att en viss referens finns med i referenslistan, vad det är för typ av dokument och hur den använts i texten.

Detta för oss in på **Martyns** forskning och hans skepticism gentemot Kesslers teori. Martyn menar att en bibliografisk koppling endast är en indikation av okänt värde på att dokumenten med gemensamma referenser har ett samband.<sup>101</sup> Den här åsikten får stöd i den här uppsatsen. Tar man initialdokument 10 som exempel kan man klart se att trots en bibliografisk koppling så har de relaterade dokumenten inget samband med informationsbehovet man skapat utifrån initialdokumentet. Sökningen ger därför 0 % i resultat genom FRR.

---

<sup>100</sup> Kessler (1962), s. 11.

<sup>101</sup> Martyn (1964), s. 236.

**Weinberg** kritiserade också Kessler i det att hon ansåg att hans forskning höll sig inom ett specifikt ämnesområde där dokumenten redan antas ha en koppling.<sup>102</sup> Kesslers undersökning gjordes på 8521 artiklar ur tidskriften *Physical Review*. Den här uppsatsens empiriska studie har till skillnad mot Kesslers studie utförts i en tvärvetenskaplig miljö, ISI Web of Science, där SCI ingår, något som Weinberg förespråkade för en studie. Det är svårare att uppnå en hög relevans vad gäller bibliografisk koppling i en studie som den i den här uppsatsen, som utförts inom ett tvärvetenskapligt ämne och utifrån fastställda informationsbehov. Studien är inte i den storlek som Weinberg hade föredragit men ger ändå en fingervisning om att bibliografisk koppling som sökmetod fungerar bäst inom ett specifikt ämnesområde som Kesslers studie visade.<sup>103</sup>

**Sen** och **Gan** utvecklade idén om bibliografisk koppling med avseende på kopplingsstyrka. Om man har en stark kopplingsstyrka så är chansen stor att dokumenten behandlar samma ämnesområde.<sup>104</sup> Men, om informationsbehovet är smalt är det inte säkert att en stark kopplingsstyrka betyder att dokumenten som har gemensamma referenser med initialdokumentet är relevanta med avseende på informationsbehovet som initialdokumentet skapat. Det är också viktigt att fundera på vad som menas med stark kopplingsstyrka. Ibland kan en kopplingsstyrka på 4 ett relativt lågt tal, visa på stor samhörighet mellan dokument. De gemensamma referenserna bör ibland också ställas mot den totala summan av referenser som ett dokument använt sig av. Delar dokumenten 4 referenser av 10 så är kopplingen förmodligen större än om de delar 4 referenser av 83. Just denna aspekt har inte undersökts i den här uppsatsen men kan vara ett intressant uppslag till vidare forskning.

**Vladutz** och **Cook** utförde en stor tvärvetenskaplig studie i SCI, en studie som liknar den som Weinberg efterlyste. I studien fann de att bibliografisk koppling är ett effektivt sätt att återvinna ämnesrelaterade dokument med.<sup>105</sup> I den här uppsatsens empiriska studie har inte ämnesrelevans varit det enda kriteriet. Specifika informationsbehov har varit grunden för relevanskriterierna. Frågan man kan ställa sig är om man som användare på ett bibliotek oftare har ett intresse av att få information om en hel ämnesdisciplin eller information om det specifika informationsbehovet.

Vladutz och Cook fick ett högt resultat, 86 %, i sin utvärdering av ämnesrelevans. Denna höga siffra ska dock sättas i relation till det faktum att de använder sig av enbart en databas och trots att hela databasens innehåll ingick i studien så var den inte lika omfattande 1981 som idag. I den här uppsatsens empiriska studie har SCI samt två andra databaser ingått. Dessutom har tidsspannet varit mellan 1986 och 2006, alltså tjugo år. Den här empiriska studien har alltså utförts i en större samling, med en större kvantitet av dokument.

**Peters, Braam** och **van Raan** kom i sin forskning fram till att dokument med en bibliografisk koppling har en betydligt större ämnesrelation än andra dokument. Deras forskning visar också på att dokument som delar en citering till en flitigt citerad

---

<sup>102</sup> Weinberg (1974), s. 195.

<sup>103</sup> Kessler (1962), s. 11.

<sup>104</sup> Sen & Gan (1983), s. 78-79.

<sup>105</sup> Vladutz & Cook (1984), s. 204.

publikation har en högre ämnesrelation än dokument som saknar denna koppling.<sup>106</sup> Detta kan man ibland se i den här uppsatsen, särskilt i vissa enstaka fall där man ser att ett visst dokument återkommer i flera dokumentets referenslistor. Ett exempel på detta kunde ses vid redan nämnda initialdokument 2 där 6 av 7 relevanta dokument bland annat delar två referenser som behandlar samma ämnesområde, ett område som kan kopplas till B&I.

Att dela på ett flitigt citerat dokument behöver dock inte alltid innebära att de kopplade dokumenten har en ämnesmässig relevans. Vid initialdokument 4 till exempel ser man ingen ämnesmässig relation mellan initialdokumentet och de bibliografiskt kopplade dokumenten trots att alla dokument i träfflistan delar exakt samma två referenser med initialdokumentet. Referenserna har bevisligen en koppling till initialdokumentet men inte av det slag att kopplingen återvinner dokument inom exakt samma ämnesområde.

---

<sup>106</sup> Peters, Braam & van Raan (1995), s. 19-21.

## 7. Slutsatser

Med avseende på precision kan man slå fast att sökmetoden FRR uppvisar ett bättre resultat än AK både när det gäller uträkningar med effektivitetsmättet P(10) och AP, 30,50% och 17,50% respektive 38,93% och 16,39%. I enstaka fall presterar dock AK-metoden bättre än FRR. Eftersom medelvärdet för FRR är större än medelvärdet för AK i både P(10) och AP kan man dock konstatera att sökning genom bibliografisk koppling i den här uppsatsen oftast är bättre än sökning med söktermer då dessa tenderar att bli för smala eller för breda. Ingen metod visar sig dock vara överlägset bättre än den andra.

Genom att testa metoderna i ISI Web of Science med avseende på precision kan för- och nackdelar med att utgå från ett initialdokument och använda sig av pearl growing-metoden jämföras och utvärderas. Båda sökmetoderna som använts i den empiriska studien för detta syfte, FRR och AK, har visat både positiva och negativa sidor. Metodernas för- och nackdelar kan till viss del förklara de skillnader som föreligger när det kommer till skillnader i precision.

Fördelar med FRR är bland annat att det är en automatisk tjänst, som på så sätt är tidsbesparande, något som gynnar både bibliotekarier och användare. En annan fördel är att genom att hänvisas till dokument med åtminstone en likadan referens som initialdokumentet kan användaren komma i kontakt med dokument som kanske inte fullt ut behandlar samma ämne som initialdokumentet men som kan vara av intresse i alla fall. Slutligen kan det nämnas att FRR kan vara ett bra komplement till andra sökmetoder, sökmetoder där söktermer används.

FRR-metoden har dock också sina nackdelar. En nackdel är själva grundprincipen med metoden. FRR-metoden återvinner dokument med minst en gemensam referens men en kopplingsstyrka på 1 är inte alltid speciellt bra för relevansen. Bibliografiskt kopplade dokument kan behandla vitt skilda ämnen varpå de inte kan anses tillfredsställa informationsbehovet. Detta påverkar relevansbedömningen och i slutändan resultatet.

Fördelar med AK-metoden är att förslag till söktermer ges av dokumentets författare, vilket är tidsbesparande och kanske ger uppslag till söktermer som användaren själv inte hade tänkt på. Det är en metod som relativt enkelt skulle kunna implementeras som automatisk tjänst och på så sätt spara tid för användarna.

Nackdelar med AK-metoden är att den är beroende av att nyckelorden vare sig är alltför breda eller alltför smala. För breda söktermer kan leda till dålig precision och för smala söktermer till dålig recall. Vid en automatisering av tjänsten är det dessutom av stor vikt att nyckelorden står i en genomtänkt ordning med det mest specifika nyckelordet först. Som det ser ut idag är det dock så att inte alla bibliografiska poster är försedda med Author Keywords. Kanske anses det som en olägenhet att lägga ner tid på att skriva in nyckelord.

Intentionen med den här granskningen har varit att se om pearl growing är en effektiv och meningsfull metod, framförallt vid sökning efter vetenskapliga artiklar inom ett specifikt ämnesområde. Om 100 % i effektivitet anses som det optimala resultatet så uppvisade P(10) och AP relativt låga siffror i medelvärde. Sökmetoderna leder dock till

att nya relevanta dokument hittas i de flesta fallen, vilket gör att de kan anses vara meningsfulla metoder.

Klassisk pearl growing som sökmetod är generellt sett effektiv om man har ett bra initialdokument att utgå ifrån. Genom att använda nyckelord från titel och abstrakt och eventuell fulltext underlättas sökningen eftersom användaren på så sätt får tips på söktermer. Pearl growing är en bra metod både om man är väl insatt i det ämne man vill hitta mer information om och om man inte besitter några förkunskaper alls. I båda fallen kan metoden bidra med nya uppslag.

Emellertid kan pearl growing vara en krävande sökmetod. Den förutsätter att användaren har ett initialdokument att utgå ifrån. Dessutom kan det ibland vara svårt att avgöra när sökningen ska avslutas eftersom det är en sökprocess som kan fortgå länge. Tar man sökmetoden FRR som exempel så kopplas initialdokumentet, pärlan, till ett relevant dokument som den har en gemensam referens med. Här skulle en användare kunna fortsätta och använda pearl growing-metoden FRR på det nya dokumentet och så vidare tills informationsbehovet är tillfredsställt. En sökning genom AK-metoden kan också vara tidskrävande och pågå så länge som användaren finner uppslag på nya söktermer.

I den här uppsatsen har alltså två pearl growing-metoder testats för att titta på hur de fungerar när man utgår ifrån specifika informationsbehov. Man hade kunnat göra så att man nöjt sig med att titta efter om dokumenten rent generellt faller inom ämnesområdet som valts till den här uppsatsen, B&I. Sådana studier har dock redan gjorts, inom bibliografisk koppling, och då har det konstaterats att bibliografisk koppling fungerar. Hade den här uppsatsen bara utgått från en generell koppling mellan initialdokument och andra dokument hade resultatet antagligen blivit avsevärt bättre. Dock är det sällan realistiskt att studenter/forskare är intresserade av ett helt forskningsfält utan har mer avgränsade informationsbehov, som dock ibland kan utökas genom pearl growing.

Utifrån den här uppsatsens empiriska studie kan inga generella slutsatser dras vad gäller vilken av sökmetoderna FRR och AK som är bäst. Tittar man på P(10) och AP visar FRR på ett bättre resultat än AK men resultatet är inte statistiskt säkerställt. Dock kan man säga att metoderna kompletterar varandra.

## 8. Sammanfattning

Databaser är en integrerad del av arbetet på ett bibliotek varpå det är av stor vikt att känna till olika sökmetoder, allt för att få en så effektiv informationsåtervinning som möjligt. Den här uppsatsen har undersökt två sökmetoder i ISI Web of Science, en tvärvetenskaplig databasportal. Sökmetoderna har pearl growing-karaktär, med vilket man menar att de utgår från ett initialdokument, en pärla, för att sedan försöka återvinna liknande dokument. Den ena metoden är Find Related Records, FRR, som återvinner dokument genom bibliografisk koppling och är en automatisk tjänst i ISI Web of Science. Den andra metoden utgår från de två första nyckelorden ur Author Keywords-fältet, AK, ett fält som ingår i den fullständiga bibliografiska posten i ISI Web of Science.

Syftet med den här uppsatsen har varit att testa och jämföra två pearl growing-metoder för att se vilken utav dem som uppvisar bäst återvinningseffektivitet vad gäller specifika informationsbehov inom ämnet biblioteks- och informationsvetenskap.

Uppsatsen har utgått från följande frågeställningar:

- Vilken precision uppvisar de olika metoderna?
- Vad beror eventuella skillnader i precision på?
- På vilket sätt är pearl growing en effektiv och meningsfull metod, framförallt vid sökning efter vetenskapliga artiklar inom ett specifikt ämnesområde?

För att sätta in den empiriska studien i ett vidare sammanhang innehåller uppsatsen en teoretisk bakgrund av IR i kapitel 2. Här presenteras IR och de klassiska IR-modellerna. Därpå följer ett avsnitt om utvärdering av IR-system där Cranfield och TREC varit betydande eftersom Cranfield till exempel gett upphov till effektivitetsmåten precision och recall, två effektivitetsmått som definieras i avsnittet därefter. Relevansbegreppet problematiseras i efterföljande avsnitt eftersom det är ett mångtydigt begrepp som dock är av stor vikt inom IR och för den här uppsatsens empiriska studie. Den teoretiska bakgrunden innehåller även en presentation av pearl growing och bibliografisk koppling.

Kapitel 3 behandlar tidigare forskning inom bibliografisk koppling, forskning som sedan diskuteras i relation till uppsatsen i kapitel 6.

Kapitel 4 är ett metodkapitel. Där beskrivs ISI Web of Science, den miljö där den empiriska studien utfördes. Metoden utgick från 20 initialdokument som återvanns i ISI Web of Science. Utifrån initialdokumenten söktes likartade dokument genom sökmetoderna FRR och AK enligt vissa fastställda kriterier. Träfflistornas 20 första dokument relevansbedömdes, enligt en binär relevansskala av uppsatsens författare, utifrån relevanskriterier som utgick från informationsbehov för varje initialdokument. Sökmetodernas effektivitet beräknades genom måtten Precision vid DCV=10, P(10), samt Uninterpolated Average Precision, AP. Eftersom några dokument inte gick att relevansbedöma ersattes de av näst följande dokument i träfflistan. För att kontrollera om dessa rockader påverkade resultatet genomfördes en parallell uträkning där Bästa/Sämsta tänkbara utfall beräknades. Resultatet av den empiriska studien redovisas i form av tabeller och diagram i kapitel 5.



I kapitel 6 diskuteras resultatet som visar att FRR-metoden presterar bättre än AK-metoden både vad gäller P(10) och AP. Vid enstaka fall presterar dock AK-metoden bättre än FRR-metoden. Denna inkonsekvens kan ha bidragit till att det inte gick att statistiskt säkerställa resultatet. Ingen av sökmetoderna kan alltså sägas vara överlägset bättre än den andra. Sökmetodernas olika grundprinciper gör dock att de kompletterar varandra bra. Uträkningen av Bästa/Sämsta tänkbara utfall visade att resultatet inte påverkades nämnvärt av att inte alla dokument gick att relevansgranska. Kapitlet avslutas med ett avsnitt där den tidigare forskningen i bibliografisk koppling diskuteras i förhållande till uppsatsens resultat.

Avslutningsvis följer kapitel 7 där uppsatsens frågeställningar besvaras och allmänna slutsatser dras. Några för- och nackdelar med de två sökmetoderna diskuteras för att försöka ge en förklaring till de skillnader som föreligger i precision dem emellan. Trots att pearl growing-metoderna som jämförs i den här uppsatsen uppvisar relativt låga medelvärden så leder de ändå till att nya relevanta dokument hittas. Pearl growing kan därför slutligen anses vara en förhållandevis bra sökmetod som med fördel kan kombineras med andra sökmetoder.

## Referenser

Baeza-Yates, Ricardo & Ribeiro-Neto, Berthier (1999). *Modern Information Retrieval*. Harlow: Addison-Wesley.

Chowdhury, G.G. (1999). *Introduction to Modern Information Retrieval*. London: Library Association Publishing.

Efthimiadis, Efthimis N. (1996). Query Expansion. Ingår i Williams, Martha E. ed. *Annual Review of Information Science and Technology: Vol. 31*. Medford, NJ: Information Today, Inc. 1996. s. 121-187.

Gravetter, Frederick J. & Wallnau, Larry B. (2000). *Statistics for the Behavioral Sciences*, 5 ed. Belmont, CA: Wadsworth/Thomson Learning.

Harter, Stephen P. (1986). *Online Information Retrieval: Concepts, Principles, and Techniques*. Orlando: Academic Press.

Jarneving, Bo (2005). *The Combined Application of Bibliographic Coupling and the Complete Link Cluster Method in Bibliometric Science Mapping*. Borås: Valfrid.

Kagolovsky, Yuri & Moehr, Jochen R. (2003). Current Status of the Evaluation of Information Retrieval. *Journal of Medical Systems*, vol. 27, no. 5, s. 409-424.

Kessler, M. M. (1963). Bibliographic Coupling Between Scientific Papers. *American Documentation*, 14, s. 10-25.

Kärki, Riita & Kortelainen, Terttu (1998). *Introduktion till bibliometri*. Helsingfors: Nordinfo.

Körner, Svante & Wahlgren, Lars (2000). *Statistisk dataanalys*. Lund: Studentlitteratur.

Large, Andrew, Tedd, Lucy A. & Hartley Richard J. (2001). *Information Seeking in the Online Age: Principles and Practice*. München: Saur.

Marchionini, Gary (1995). *Information Seeking in Electronic Environments*, Cambridge: Cambridge University Press.

Martyn, John (1964). Bibliographic Coupling. *Journal of Documentation*, vol. 20, s. 236.

Mizzaro, Stefano (1998). How Many Relevances in Information Retrieval?. *Interacting with Computers*, vol. 10, no. 3, s. 303-320.

Peters H.P.F., Braam, R.R. & van Raan, A.F.J. (1995). Cognitive Resemblance and Citation Relations in Chemical Engineering Publications. *Journal of the American Society for Information Science*, vol. 46, no. 1, s. 9-21.

Ramer, Sheryl L. (2005). Site-ation Pearl Growing: Methods and Librarianship History and Theory. *Journal of the Medical Library Association*, vol. 93, no. 3, s. 397-400.

Reitz, Joan M. (2004). *Dictionary for Library and Information Science*. Westport, Conn: Libraries Unlimited.

Saracevic, Tefko (1975). Relevance: A Review of and a Framework for the Thinking on the Notion in Information Science. *Journal of the American Society for Information Science*, vol. 26, s. 321-343.

Sen, Subir K. & Gan, Shymal K. (1983). A Mathematical Extension of the Idea of Bibliographic Coupling and its Applications. *Annals of Library Science and Documentation*, vol. 30, no. 2, s. 78-82.

Tenopir, Carol (2001). The Power of Citation Searching. *Library Journal*, vol. 126, no. 18, s. 39-40.

Text Retrieval Conference (TREC) (2004-01-30). Overview.  
<http://trec.nist.gov/overview.html> [2006-02-17]

Vladutz, George & Cook, James (1984). Bibliographic Coupling and Subject Relatedness. *Proceedings of the 47<sup>th</sup> ASIS Annual Meeting*, vol. 21, s. 204-207.

Web of Science – Thomson Scientific. <http://scientific.thomson.com/products/wos/> [2006-03-29].

Weinberg, Bella Hass (1974). Bibliographic Coupling: a Review. *Information Storage and Retrieval*, vol. 10, no. 5/6, s. 189-196.

White, Howard D. & McCain, Katherine W. (1998). Visualizing a Discipline: An Author Co-Citation Analysis of Information Science, 1972-1995. *Journal of the American Society for Information Science*, vol. 49, no. 4, s. 327-355.

## Bilaga 1. Initialdokumenten

**1. Informationsbehov:** Forskning om biblioteks- och informationsvetenskap.

**Relevans:** Dokumenten kan innehålla generell information om forskning om B&I, diskussioner gällande forskningsfältets ramar/förutsättningar/begränsningar.

**Title:** Philosophical foundations and research relevance: issues for information research

**Abstract:** This paper examines three ideas that affect the nature of research in information science: (1) the fact of the incoherent nature of information science, which results from the concept of 'information' being dealt with at different integrative levels; (2) the lack of an over-arching philosophical framework that might guide the development of methods; and (3) the problem of grounding research in the reality of everyday professional practice. It is suggested that, at one level, phenomenology offers an integrative philosophical perspective that might also help to resolve the research/practice split.

**Author Keywords:** information science; librarianship; research; philosophical framework; methodologies; models; relevance; professional practice; knowledge organization; integrative levels; phenomenology

**2. Informationsbehov:** Genetiska algoritmer inom informationsåtervinning.

**Relevans:** Dokumentet bör innehålla information om genetiska algoritmer i samband med informationsåtervinning.

**Title:** Boolean queries optimization by genetic algorithms

**Abstract:** Information retrieval systems depend on Boolean queries. Proposed evolution of Boolean queries should increase the performance of the information retrieval system. Information retrieval systems quality are measured in terms of two different criteria, precision and recall. Evolutionary techniques are widely applied for optimization tasks in different areas including the area of information retrieval systems. In information retrieval applications both criteria have been combined in a single scalar fitness function by means of a weighting scheme 'harmonic mean'. Usage of genetic algorithms in the Information retrieval, especially in optimizing a Boolean query, is presented in this paper. Influence of both criteria, precision and recall, on quality improvement are discussed as well.

**Author Keywords:** evolutionary algorithms; genetic algorithms; genetic programming; information retrieval; Boolean query

**3. Informationsbehov:** Anonymitet i samband med kommunikation via nätet.

**Relevans:** Dokument ska informera om eller problematisera anonymitet vid kommunikation på nätet.

**Title:** Identification of comment authorship in anonymous group support systems

**Abstract:** This study examines whether technically "anonymous" comments entered by participants during group support system (GSS) brainstorming sessions are, in fact, unidentifiable. Hypotheses are developed and tested about the influences of comment length, comment evaluative tone, duration of group membership, and prior communication among group members on the accuracy of attributions they made about the identity of the authors of these technically anonymous comments. Data on prior communication and group history about each of the 32 small groups was collected before participants began using a GSS for brainstorming. Immediately after the session, each member was asked to attribute authorship to a sample of the session's anonymous comments (comment authorship was known to the researchers). The study's participants made attributions that were significantly more accurate than chance guessing. Factors that had a positive influence on attribution accuracy include evaluative tone of comments (especially humorous comments) and amount of prior communication received from other group members. Vividness of comment tone and comment length was not significantly correlated with attribution accuracy. Although the attributions of anonymous comments were more accurate than

expected by chance, most of the attributions were incorrect. Implications and consequences of both accurate and inaccurate attribution are discussed along with suggestions for future research.

**Author Keywords:** anonymity; computer-mediated communication; group support systems; social networks

**4. Informationsbehov:** Klassifikation i samband med informationsåtervinning.

**Relevans:** Alla dokument som behandlar klassificering i samband med informationsåtervinning.

**Title:** Some thoughts on classification for retrieval

**Abstract:** Purpose-This paper, originally published in 1970, considered the suggestion that classifications for retrieval should be constructed automatically and raised some serious problems concerning the sorts of classification which were required, and the way in which formal classification theories should be exploited, given that a retrieval classification is required for a purpose. These difficulties had not been sufficiently considered, and the paper, therefore, aims to attempt an analysis of them, though no solutions of immediate application could be suggested.

Design/methodology/approach-Starting with the illustrative proposition that a polythetic, multiple, unordered classification is required in automatic thesaurus construction, this is considered in the context of classification in general, where eight sorts of classification can be distinguished, each covering a range of class definitions and class-finding algorithms.

Findings-Since there is generally no natural or best classification of a set of objects as such, the evaluation of alternative classifications requires either formal criteria of goodness of fit, or, if a classification is required for a purpose, a precise statement of that purpose. In any case a substantive theory of classification is needed, which does not exist; and, since sufficiently precise specifications of retrieval requirements are also lacking, the only currently available approach to automatic classification experiments for information retrieval is to do enough of them.

Originality/value-Gives insights into the classification of material for information retrieval.

**Author Keywords:** information retrieval; classification

**5. Informationsbehov:** Artificial Intelligence (AI) i förhållande till IR och/eller dokumentåtervinning med hjälp av den probabilistiska modellen.

**Relevans:** Artificial Intelligence (AI) i förhållande till IR och/eller dokumentåtervinning med hjälp av den probabilistiska modellen.

**Title:** Information retrieval and artificial intelligence

**Abstract:** This paper addresses the relations between information retrieval (IR) and AI. It examines document retrieval, summarising its essential features and illustrating the state of its art by presenting one probabilistic model in detail, with some test results showing its value. The paper then analyses this model and related successful approaches, concentrating on and justifying their use of weak, redundant representation and reasoning. It goes on to other information management tasks and considers how the concepts and methods developed for retrieval may be applied to these, concluding by arguing that such ways of dealing with information may also have wider relevance to AI. (C) 1999 Published by Elsevier Science B.V. All rights reserved.

**Author Keywords:** information retrieval; probabilistic model; Artificial Intelligence

**6. Informationsbehov:** Dokument om sökstrategier för informationsåtervinning.

**Relevans:** Dokument om sökstrategier för informationsåtervinning.

**Title:** General and efficient strategies for information retrieval

**Abstract:** To address the inefficient use of information retrieval (IR) systems such as search engines and library catalogues, we present a unified framework of strategies for information retrieval. This framework (1) contains a small set of general and efficient information retrieval strategies that are useful across many IR systems, and (2) can be used to identify key missing functionality in IR systems, and to design training approaches that lead to the efficient retrieval of information.

**Author Keywords:** strategies; information retrieval; search; efficient

**7. Informationsbehov:** Referenstjänster och digitala bibliotek.

**Relevans:** Dokument som behandlar digitala bibliotek och deras nätbaserade referenstjänster.

**Title:** Digital libraries and reference services: present and future

**Abstract:** Reference services have taken a central place in library and information services. They are also regarded as personalised services since in most cases a personal discussion takes place between a user and a reference librarian. Based on this, the librarian points to the sources that are considered to be most appropriate to meet the specific information need(s) of the user. Since the Web and digital libraries are meant for providing direct access to information sources and services without the intervention of human intermediaries, the pertinent question that appears is whether we need reference services in digital libraries, and, if so, how best to offer such services. Current digital libraries focus more on access to, and retrieval of, digital information, and hardly lay emphasis on the service aspects. This may have been caused by the narrower definitions of digital libraries formulated by digital library researchers. This paper looks at the current state of research in personalised information services in digital libraries. It first analyses some representative definitions of digital libraries in order to establish the need for personalised services. It then provides a brief overview of the various online reference and information services currently available on the Web. The paper also briefly reviews digital library research that specifically focuses on the personalisation of digital libraries and the provision of digital reference and information services. Finally, the paper proposes some new areas of research that may be undertaken to improve the provision of personalised information services in digital libraries.

**Author Keywords:** libraries; information services; information technology

**8. Informationsbehov:** Klassifikationsteori och/eller klassifikation i samband med informationsåtervinning.

**Relevans:** Dokument rörande klassifikationsteori och/eller klassifikation i samband med informationsåtervinning.

**Title:** Revisiting classification for retrieval

**Abstract:** Purpose-This short note seeks to respond to Hjørland and Pederson's paper "A substantive theory of classification for information retrieval" which starts from Sparck Jones's, "Some thoughts on classification for retrieval", originally published in 1970.

Design/methodology/approach-The note comments on the context in which the 1970 paper was written, and on Hjørland and Pedersen's views, emphasising the need for well-grounded classification theory and application.

Findings-The note maintains that text-based, a posteriori, classification, as increasingly found in applications, is likely to be more useful, in general, than a priori classification.

Originality/value-The note elaborates on points made in a well-received earlier paper.

**Author Keywords:** information retrieval; classification

**9. Informationsbehov:** Hur ska bibliotek förhålla sig till kontroversiella ämnen i sina samlingar?

**Relevans:** Dokument rörande kontroversiella ämnen i bibliotekets bestånd. Även dokument om klassificering av kontroversiella ämnen som till exempel pornografi.

**Title:** Deep classification: pornography, bibliographic access, and academic libraries

**Abstract:** This study examines the mainstreaming of pornography in the context of current economic, popular culture, and academic trends. As pornography becomes part of popular culture, it simultaneously becomes an area of focus for academics and therefore presents particular challenges for college and university libraries. Both physically and conceptually, academic libraries must find a place for pornography on the shelves and in the array of knowledge structured by bibliographic access systems. This study looks at how the variety of issues, concepts, and genres of pornography considered in academic discourse could be accommodated within access systems by examining the way in which the adult industry itself classifies pornographic films. Specifically, the terms used by the adult industry to classify these films could be grouped within newly developed categories. The identification of the categories would not be predicated on characteristics of porn films alone. Instead, the categories would encompass specific topics, concepts, and subject areas that connect pornography to mainstream culture. Using classifications from four different adult industry sources, four sample categories are presented that could serve as a model for how pornographic concepts could be accommodated within existing bibliographic access systems. (C) 2002 Elsevier Science Ltd. All rights reserved.

**Author Keywords:** pornography; censorship; intellectual freedom; classification theory; classification systems; subject access; collection development

**10. Informationsbehov:** Utvecklingen av skolbibliotek.

**Relevans:** Dokument rörande skolbibliotekens historia och/eller utveckling.

**Title:** New developments on the Turkish school library scene

**Abstract:** The overall purpose of this article is to describe the history, growth and development of school libraries in Turkey from 1923 to 2004. For now and the foreseeable future, school librarians will be simultaneously working in the library of yesterday and deeply affected by the library of tomorrow. Changing information needs make it necessary to extend school library services to include new information resources. School librarians must help students understand their information needs and the resources and information technologies available. The research on which this article is based used the survey method. Data were collected through literature analysis, questionnaires, interviews and observation. Observations and interviews were conducted and 3000 questionnaires were distributed in 100 secondary education institutions in Ankara, Turkey during the academic year 2003-4. The research in this article explores the historical background and the current status, role and function of school libraries in providing information resources to help meet the information needs of students in Turkey. It is concluded that in order to optimally deliver information services in secondary education institutions, it is necessary to connect with and guide users by all means available, thereby providing endless possibilities for perpetual connectivity and human development.

**Author Keywords:** information needs; information resources; school libraries; Turkey

**11. Informationsbehov:** Information om Filtering-system (IF-system).

**Relevans:** Dokument rörande Information Filtering-system och/eller profilbaserad information och dess effektivitet.

**Title:** Using the information structure model to compare profile-based information

**Abstract:** In the IR field it is clear that the value of a system depends on the cost and benefit profiles of its users. It would seem obvious that different users would prefer different systems. In the TREC-9 filtering track, systems are evaluated by a utility measure specifying a given cost and benefit. However, in the study of decision systems it is known that, in some cases, one system may be unconditionally better than another. In this paper we employ a decision theoretic approach to find conditions under which an Information Filtering (IF) system is unconditionally superior to another for all users regardless of their cost and benefit profiles.

It is well known that if two IF systems have equal precision the system with better recall will be preferred by all users. Similarly, with equal recall, better precision is universally preferred. We confirm these known results and discover an unexpected dominance relation in which a system with lower recall will be universally preferred provided its precision is sufficiently higher.

**Author Keywords:** performance; evaluation; formal model; user-profile; information economics

**12. Informationsbehov:** Läsvanor hos forskare.

**Relevans:** Dokument rörande läsvanor hos forskare. Här ingår sökning, återvinning, läsning och användning av information.

**Title:** How, what, and why science faculty read

**Abstract:** One part of a library use survey administered in 1993 to the users of the libraries of the University of Tennessee at Knoxville was analyzed to provide a profile of the reading behaviors of science faculty. The survey questions were grouped into the four steps in accomplishing a document reading: finding, getting, reading, and using a document. The science faculty tended to find their document by browsing. The document was generally obtained from a personalized subscription. Science faculty predominately read journals and books and read almost 50 articles a year from the journal they obtained their last read document from. The science faculty were mostly likely to use their reading for research. Science faculty appreciated the capabilities of electronic document transmission, but had a marked preference for paper transmission. Few differences were found between science and non-science faculty, mostly in their specific concerns for electronic or print transmission and in the uses of their readings.

**Author Keywords:** scientists; faculty; reading behaviors; browsing; document transmission

**13. Informationsbehov:** Information om Impact factor.

**Relevans:** Dokumenten ska behandla fenomenet Impact factor.

**Title:** Use of Journal Citation Reports and Journal Performance Indicators in measuring short and long term journal impact

**Abstract:** The impact factor has become the subject of widespread controversy. It has gradually developed to mean both journal and author impact. The emphasis on impact factors obscures the main purpose of bibliographic databases created at the Institute for Scientific Information. I will here show how two of these databases, Journal Citation Reports and the Journal Performance Indicators, can be used to study scientific journals and the articles they publish, as well as the evolution of scientific fields.

**Author Keywords:** bibliometrics; citation analysis; impact factor; journal article; library science; medical informatics; medical literature analysis and retrieval system.

**14. Informationsbehov:** Referensservice på nätet.

**Relevans:** Dokumenten ska ge information om attityder gentemot biblioteksreferenstjänster på nätet, alternativt utvärderingar av sådana tjänster.

**Title:** An evaluative case study of a real-time online reference service

**Abstract:** This case study evaluated a real-time online interactive reference service. Three data gathering methods were used. First, general usage statistics were generated electronically from the service provider. Second, patron evaluations were collected electronically. Third, questionnaires were mailed to the librarians who provided the service. The general usage statistics revealed an overall pattern of growth in the number of participating libraries, the number of patrons, the number of questions and the total time spent answering them in minutes. Patron evaluations of the service were positive, most of the patrons' questions were answered. They all liked the service and intended to use it again. The librarians' responses were also positive. Most of the librarians were able to find quality Internet Web sites containing the answers to their patrons' questions and when they could not, they often found the answers in a print resource. This finding suggested that regardless of the setting, print resources still had an important role to play in the provision of reference service.

**Author Keywords:** reference services; information; libraries; services



**15. Informationsbehov:** Förändringar och omstruktureringar av bibliotek i världen under de senaste två decennierna och inför framtiden.

**Relevans:** Förändringar och omstruktureringar av bibliotek i världen under de senaste två decennierna och inför framtiden.

**Title:** Restructuring library organizations for the twenty-first century: the future of user-oriented services in Korean academic libraries

**Abstract:** This article examines why Korean academic libraries embarked on restructuring their organization and identifies possible solutions to some problems that these libraries face. The conclusion is drawn from the results of a case study on Sejong University Library that Korean academic librarians should become user-centered professionals. After restructuring, the user service librarian's job content has changed to concentrate on servicing users on specific subjects. The subject librarian will take care of reference services, SDI, DB instruction and book selection on the same subject.

**Author Keywords:** restructuring; organization; library users; acquisitions (library materials); South Korea; academic libraries

**16. Informationsbehov:** Hur upplever män att arbeta inom kvinnodominerade yrken, till exempel bibliotek?

**Relevans:** Dokument som behandlar män som arbetar inom kvinnodominerade yrken, till exempel på bibliotek.

**Title:** Masculinity at work: the experiences of men in female dominated occupations

**Abstract:** This article presents the findings of a research project on the implications of men's non-traditional career choices for their experiences within the organization and for gender identity. The research is based on 40 in-depth interviews with male workers from four occupational groups: librarians, cabin crew, nurses, and primary school teachers. Results suggest a typology of male workers in female dominated occupations: seekers (who actively seek the career): finders (who find the occupation in the process of making general career decisions) and settlers (who settle into the career after periods of time in mainly male dominated occupations). Men benefit from their minority status through assumptions of enhanced leadership (the assumed authority effect), by being given differential treatment (the special consideration effect) and being associated with a more careerist attitude to work (the career effect). At the same time, they feel comfortable working with women (the zone of comfort effect). Despite this comfort, men adopt a variety of strategies to re-establish a masculinity that has been undermined by the feminine nature of their work. These include re-labelling, status enhancement and distancing from the feminine. The dynamics of maintaining and reproducing masculinities within the non-traditional work setting are discussed in the light of recent theorizing around gender, masculinity and work.

**Author Keywords:** gender; masculinity; non-traditional occupations; post-structuralism

**17. Informationsbehov:** Indexering av skönlitteratur

**Relevans:** Dokument som handlar om indexering av skönlitteratur.

**Title:** Consistency of subject indexing of novels by public library professionals and patrons

**Abstract:** The paper discusses the consistency of fiction indexing of library professionals and patrons based on an empirical test. Indexing was carried out with a Finnish fictional thesaurus and all of the test persons indexed the same five novels. The consistency of indexing was determined to be low; several reasons are postulated. Also an algorithm for typified indexing of fiction is given as well as some suggestions for the development of fiction information retrieval systems and content representation.

**Author Keywords:** indexes; information retrieval

**18. Informationsbehov:** Olika sätt att indexera på för att underlätta dokumentåtervinning.

**Relevans:** Olika sätt att indexera på för att underlätta dokumentåtervinning.

**Title:** Using citations to facilitate precise indexing and automatic index creation in collections of research papers

**Abstract:** We describe Rosetta, a digital library system for scientific literature. Rosetta makes it easy for people to find the information for which they are looking even when using short, imprecise queries. Rosetta indexes research articles based on the way they have been described when cited in other documents. The concise descriptions that occur in citations are similar to the short queries people typically form when searching; therefore, citations may make a better basis for indexing than do the words used within a research article itself. Using this indexing technique we are able to provide a user interface that presents users with an automatically generated directory of the information space surrounding a query. Our objective with this interface is to present people with the information for which they have asked as well as the information for which they may have intended to ask. (C) 2001 Elsevier Science B.V. All rights reserved.

**Author Keywords:** search engines; information retrieval; reference directed indexing; citation indexing

**19. Informationsbehov:** Internet som informationskälla

**Relevans:** Dokument som problematiserar Internet som informationskälla.

**Title:** Mozart versus Minsky: information bias on the Internet

**Abstract:** Notes that until the advent of the Internet, major libraries and text repositories were considered by all as bastions of information, but the Internet has opened up a treasure-trove of information for researchers and seekers of knowledge throughout the world. It has become common practice by researchers of all kinds to turn to the Internet as a convenient source of information. Asks how effectively does the Internet really present these researchers with a representative picture of the state of human knowledge! Identifies the potential misuse of the Internet as a source of biased information. Defines biased information as information not representative of the state of human knowledge. In order to provide a basis for comparison, begins by defining the concept of a representative subset, which is used as a basis for comparing the nature of different information resources. This is followed by a presentation of the use of the knowledge-comparison function in evaluating different information resources. Then discusses the significance of information equivalence in the information age as a basis for setting concrete goals for Internet research policy. Concludes by presenting a number of courses of action including a description of X-DEX, an Internet indexing standard based on the US Library of Congress cross-reference system.

**Author Keywords:** information; Internet; knowledge transfer; libraries

**20. Informationsbehov:** Hur konstruerar man ett digitalt bibliotek?

**Relevans:** Dokument om hur man konstruerar ett digitalt bibliotek.

**Title:** Making a digital library: The contents of the CORE project

**Abstract:** The CORE (Chemical Online Retrieval Experiment) project is a library of primary journal articles in chemistry. Any library has an inside and an outside; in this article we describe the inside of the library and the methods for building the system and accumulating the database. A later article will describe the outside (user experiences). Among electronic-library projects, the CORE project is unusual in that it has both ASCII derived from typesetting and image data for all its pages, and among experimental electronic-library projects, it is unusually large. We describe here (a) the processes of scanning and analyzing about 400,000 pages of primary journal material, (b) the conversion of a similar amount of textual database material, (c) the linking of these two data sources, and (d) the indexing of the text material.

**Author Keywords:** algorithms; design; experimentation; human factors; image segmentation