

Scopus doorgelicht en vergeleken

De dekking en functionaliteit van de citatiedatabase Scopus,
inclusief vergelijkingen met Web of Science en Google Scholar



Jeroen Bosman
Ineke van Mourik
Menno Rasch
Eric Sieverts
Huib Verhoeff

Juni 2006
© Universiteitsbibliotheek Utrecht

Scopus doorgelicht en vergeleken: de dekking en functionaliteit van de citatiedatabase Scopus, inclusief vergelijkingen met Web of Science en Google Scholar

Inhoud

1	Samenvatting	4
2	Inleiding	6
3	Verantwoording aanpak	7
4	Dekking: de inhoud van de databases	8
4.1	Het aantal ontsloten documenten	8
4.1.1	Totaal aantal documenten.....	8
4.1.2	Aantal ontsloten tijdschriften	8
4.1.3	Ontsloten documenten naar documenttype	10
4.1.4	Ontsloten documenten en tijdschriften naar vakgebied	11
4.2	Dekking op basis van voorbeeldzoekacties.....	17
4.3	Dekkingsperiode.....	18
4.4	Actualiteit	22
4.5	Aard van de opgenomen gegevens per document.....	24
4.6	Citatiegegevens: de dekking van “citerende artikelen”.....	26
5	Zoekfunctionaliteit, interface, snelheid en gebruiksgemak.....	30
5.1	Zoekfunctionaliteit	30
5.2	Interface, snelheid en gebruiksgemak	31
5.3	Vakgebiedsindeling.....	31
6	Beoordeling door gebruikers.....	34
6.1	Interviews met veelgebruikende onderzoekers	34
6.2	Webenquete onder gebruikers.....	35
6.2.1	Conclusies	35
6.2.2	Doel van het onderzoek.....	35
6.2.3	Opzet	36
6.2.4	Uitkomst digitale vragenlijst	36
7	Samenvatting vakspecifieke resultaten	41
7.1	Overzicht	41
7.2	Opmerkingen per UBU-vakgebied.....	41
7.2.1	Aardwetenschappen	42
7.2.2	Biologie	42
7.2.3	Diergeneeskunde	42
7.2.4	Economie.....	42
7.2.5	Farmacie	43
7.2.6	Geneeskunde	43
7.2.7	Godgeleerdheid	43
7.2.8	Letteren.....	43
7.2.9	Milieukunde	43
7.2.10	Natuur- en Sterrenkunde	44
7.2.11	Rechtsgeleerdheid	44
7.2.12	Scheikunde	44
7.2.13	Sociale Geografie en Planologie	44
7.2.14	Sociale Wetenschappen.....	45

7.2.15	Wijsbegeerte.....	45
7.2.16	Wiskunde en Informatica	45
Literatuur.....		47
Bijlage VI: Niet-aangehaalde literatuur over Scopus.....		48

1 Samenvatting

Aanleiding

- **Scopus is op de markt gebracht** door Elsevier; het is een product dat concurrentie brengt in een segment van de markt voor bibliografische databases waar voorheen ISI met Web of Science een monopolie had. Kostenoverwegingen, positieve eerste reacties maken een nadere verkenning zinnig.
- **Veranderende zoekgewoontes** bij studenten (minder nadruk op vakspecifieke ingangen, meer links volgend dan systematisch zoekend) en andere concurrenten (met name Google Scholar, wellicht ook Windows Live Academic) maken het verkennen van multidisciplinaire databases belangrijk.

Aanpak

- **In de diepte**; het onderzoek kan ook als basis dienen voor latere evaluaties van andere producten en het is tevens een test voor hoe je zoiets kunt aanpakken en welke data beschikbaar zijn voor vergelijkingen.
- **Eigen onderzoek**, want er is weinig gedegen literatuur (behalve Jasco 2005 en Pipp 2006).
- Sessies met en input van vakspecialisten
- Feedbacksessie met producent
- **Gebruikersinput**: interviews en gebruikersenquête

Onderzoeksresultaat dekking

- **Omvang records, titels.** Scopus heeft bijna 28 miljoen records; het aantal records in onze versie van WoS is met 19 miljoen kleiner, maar het aantal in de complete WoS (met backfiles teruggaand tot 1945) is met 37 miljoen records groter. Scopus ontsluit ruim 15000 tijdschrifttitels tegen bijna 9000 in de WoS. Scopus ontsluit 64% van onze digitale tijdschrifttitels, tegen 53% door WoS.
- **Dekkingsperiode.** Scopus is voor 1996 5-15% kleiner, na 1996 20-45% groter dan WoS op basis van aantallen records. Voor publicaties van voor 1996 biedt Scopus nog wel een zeer onevenwichtige dekking voor de verschillende vakgebieden.
- **Documenttypen.** De totale database van Scopus bestaat voor 95% van de records uit beschrijvingen van tijdschriftartikelen. Voor de jaren vóór 1996 is het aantal niet-tijdschriftartikelen in Scopus gering, daarna stijgend tot ruim 10% in 2005. Daarmee is voor de recente jaren het aandeel niet-tijdschriftartikel een stuk groter dan in WoS (4%).
- **Vakgebiedspecifiek.** Scopus dekt uitsluitend beta- en gammadisciplines, WoS dekt daarnaast ook nog de humaniora. De dekking van Scopus is in 16 van de 18 UBU-vakgebieden 4 of meer procentpunten groter dan die van WoS op basis van aantallen tijdschrifttitels in het digitale aanbod van de UBU. De twee vakgebieden waar WoS sterker is hebben beide een alfa-signatuur. Op basis van enkele zoekacties lijkt Scopus relatief zwak voor sociologie, natuurkunde en sterrenkunde (voorzichtigheid is hier geboden en nader onderzoek nodig), maar zeer goed voor biomedische en geowetenschappen.
- **Actualiteit.** Er is, gemeten op basis van opname van afleveringen van tijdschriften en op basis van het 'vorderingspercentage' van de dekking van het lopende jaar, nauwelijks verschil tussen WoS-Scopus wat betreft snelheid waarmee nieuwe publicaties worden opgenomen.
- **Aard gegevens per record.** Scopus heeft meer trefwoorden, van auteurs maar ook vaak uit 'controlled vocabulary' (bv. MeSH), WoS heeft naast de auteurstrefwoorden

geen trefwoorden uit controlled vocabulary maar wel Keywords-plus: trefwoorden gegenereerd uit referenties.

- **Citatiegegevens.** Het verschil tussen Scopus en WoS wat betreft citatiegegevens is vrij klein, er is een sterke overlap. Telling op basis van verwijzingen naar 64 artikelen uit 1995 en 2000 wijst uit dat WoS 6% minder verwijzingen naar citerende artikelen heeft. Het verschil van deze twee met Google Scholar is groter. Weliswaar heeft Google Scholar 2% minder verwijzingen naar deze artikelen dan Scopus, het heeft daarin wel gemiddeld 5 maal meer 'unieke' verwijzende publicaties. Met name bij gamma, inclusief economie, heeft Google Scholar veel meer en meer unieke, citaties.

Onderzoeksresultaat functionaliteit

- **Verschillen in mogelijkheden.** Scopus is iets veelzijdiger en heeft met de default refine, de tabelvormige resultaten van de Citation Tracker en met de auteursidentificatie een paar duidelijke pluspunten in functionaliteit. WoS heeft net iets uitgebreidere mogelijkheden voor citatieanalyses van instellingen. NB In juni 2006 is WoS ook gekomen met een Refine tool en heeft ISI ook een auteursidentificatie aangekondigd voor WoS.
- **Snelheid.** Er is vooral groot verschil van WoS en Scopus met GS, dat vrijwel direct resultaat ophoest, en ook, afhankelijk van het type zoekactie, met de Omega-zoekmachine, die vaak ook zeer snel is. Dit kan voor gebruikers (onbewust) een belangrijk punt zijn in hun keuze voor Google Scholar. Het onderling snelheidsverschil WoS-Scopus is niet groot, maar iets in het voordeel van Scopus.

Gebruikersbeoordelingen

- **Interviews.** Veelgebruikers uit de faculteiten zijn vooral positief over helderheid van de interface Scopus en over refine en de citation tracker. De meerderheid van de geïnterviewden waardeert Scopus meer dan WoS, maar heeft tevens als harde 'eis' dat JCR beschikbaar blijft.
- **Enquête.** Uit vragen onder 81 gebruikers blijkt dat Scopus en WoS minder bekend zijn dan Google Scholar, maar de resultaten van Google Scholar worden minder goed gevonden, vooral bij AIO's/onderzoekers en dan ook weer vooral bij bètarichtingen. Scopus vind men het beste in gebruik, direct gevolgd door Google Scholar. WoS heeft hier volgens de respondenten duidelijk iets goed te maken. Ook wat betreft de relevantie van de resultaten wordt Scopus het hoogst gewaardeerd van de deze drie citatiedatabases.

2 Inleiding

In 2004 is een nieuwe multidisciplinaire database op de markt gekomen: Scopus. Hiermee is er concurrentie gekomen in een segment van de markt voor bibliografische databases waar dat er voorheen niet was (zie tabel 2.1).

Tabel 2.1 Segmenten in de markt van bibliografische bestanden

functionaliteit	multidisciplinair	vakspecifiek
titels + abstracts + volledige citatiefunctie	<ul style="list-style-type: none"> ▪ Scopus ▪ Web of Science 	-
titels + abstracts + beperkte citatiefunctie	<ul style="list-style-type: none"> ▪ EBSCO ASE ▪ Google Scholar (- abs.) ▪ Highwire (+ full text) 	<ul style="list-style-type: none"> ▪ Citeseer ▪ PsycInfo ▪ Pubmed Central ▪ RepEc Econpapers ▪ SciFinder Scholar ▪ SMEALsearch
titels + abstracts	<ul style="list-style-type: none"> ▪ DOAJ search ▪ Infotrieve articlefinder ▪ Omega (UBU) ▪ Open J-Gate 	<ul style="list-style-type: none"> ▪ (selectie:) ▪ CAB abstracts ▪ Econlit ▪ ERIC ▪ GeoArchive ▪ Geobase ▪ Georef ▪ Pubmed ▪ SocIndex ▪ Sociological Abstracts ▪ TRIS ▪ Zentralblatt MATH
titels	<ul style="list-style-type: none"> ▪ Picarta ▪ Windows Live Academic 	(voorbeeld:) <ul style="list-style-type: none"> ▪ GeoDOK

Het nieuwe product, afkomstig van Elsevier Science, is een directe concurrent van Web of Science van Thomson-ISI, dat bij de UBU in abonnement is. De eerste berichten over het product maakten duidelijk dat Scopus functioneel en qua vormgeving aantrekkelijk was, maar lieten veel vragen over de dekking van het bestand open. Een groep vak- en informatiespecialisten heeft de taak op zich genomen te onderzoeken in hoeverre Scopus daadwerkelijk als gelijkwaardig alternatief voor Web of Science gezien kan worden. Het voorliggende rapport is het verslag van dat onderzoek.

3 Verantwoording aanpak

Het onderzoek naar Scopus is uitvoerig opgezet, vanwege de hoge kosten die met het bestand en zijn concurrent zijn gemoeid en vanwege het grote belang dat aan dit soort databases wordt gehecht in het wetenschappelijk onderzoeker. Ook zijn bijna alle UU-vakgebieden erbij betrokken, waardoor het belangrijk is een goede basis te hebben voor afweging van belangen. Een andere reden om zelf Scopus uitvoerig door te lichten en te vergelijken is dat er maar weinig grondige studies bekend zijn. Alleen Jascó (2004 en 2005) en Pipp (2006) hebben Scopus en Web of Science in detail vergeleken, maar hebben dat gedaan op een moment dat Scopus nog volop in ontwikkeling was of hebben veel aspecten, zoals de dekking van vakgebieden en citatiegegevens maar beperkt onderzocht. In bibliotheekweblogs (bijvoorbeeld *One entry to research*) zijn her en der wel gedetailleerde beschouwingen te vinden, maar die gaan doorgaans maar over één klein aspect van de databases. Er is geen onderzoek bekend waarbij Scopus en WoS qua dekking worden vergeleken op het niveau van tijdschrifttitels. Ook vergelijking van dekking met vakindexen zoals uitgevoerd voor dit rapport is in de literatuur niet vindbaar.

De belangrijkste basis van dit rapport is een database van tijdschriften gedekt door Scopus, WoS en een twintig vakindexen én aanwezigheid van die tijdschriften in het digitale bezit van de UBU en de Omega-zoekmachine van de UBU. Het werk dat hiervoor is verricht is niet alleen bruikbaar voor evaluatie van Scopus en WoS maar ook voor eventuele toekomstige evaluaties van vakindexen.

Naast het dekkingsonderzoek op basis van tijdschrifttitels is ook door vele personen gedurende een half jaar de functionaliteit van Scopus getest en hebben we systematisch onderzocht hoe uitgebreid de citatieverwijzingen in Scopus zijn en hoe actueel de database is. Behalve eigen onderzoek aan de database hebben we ook naar de ervaringen en mening van gebruikers gevraagd. Daartoe is in de beginfase van het onderzoek gesproken met een twintigtal onderzoekers, allen intensieve gebruikers van Web of Science. Mede op basis van die sessies, waarin zij lieten zien hoe zij met WoS werken en wat daarbij van belang is en tevens wat hun eerste ervaringen zijn met Scopus hebben we bepaald welke aspecten in het onderzoek belicht moesten worden. Daar bleek bijvoorbeeld dat men vrijwel unaniem al overtuigd was van het gebruiksgemak van Scopus, maar nog niet van de dekking.

Naast de sessies met veelgebruikers hebben we tevens een bescheiden webenquête gehouden onder gebruikers van Scopus, waarbij ook vele studenten hun mening is gevraagd. Ondanks de uitvoerige aanpak zijn er toch nog zaken die in dit onderzoek maar terzijde worden belicht. Zo is het precieze effect van automatische stemming van Scopus, Keywords-plus van Web of Science en keywords bij Scopus nog onvoldoende helder. Verder hebben we opties voor uitvoer, attendering en personalisatie niet onderzocht. Ook de recente geïmplementeerde Author Identifier hebben we nog niet uitgebreid kunnen testen. Tenslotte is het gezien de snelle ontwikkeling van Scopus zowel als WoS van groot belang de veranderingen ook na verschijnen van dit rapport te blijven volgen en te betrekken bij eventuele aanschafkeuzes.

NB Dit rapport beschrijft onderzoek dat is uitgevoerd gericht op de behoefte en omstandigheden aan de Universiteit Utrecht. De gebruikte vakgebiedsindelingen zullen vaak niet corresponderen met die in gebruik bij andere organisaties. Ook zijn vergelijkingen gemaakt met producten, zoals de eigen Utrechtse zoekmachine Omega, die elders niet beschikbaar zijn. Informatie over licenties en bijlagen met privacygevoelige onderdelen zijn niet opgenomen in de openbare versie van dit rapport.

4 Dekking: de inhoud van de databases

Een van de belangrijkste aspecten waarop Scopus beoordeeld moet worden is de dekking, de inhoud van de database. Samen met functionaliteit en gebruiksgemak bepaalt de dekking de waarde van het bestand. In dit deel gaat het om een feitelijke inschatting van de dekking met een weging van enkele aspecten daarvan door vakspecialisten, niet om een oordeel van gebruikers. Dat doen we in hoofdstuk 6.

De dekking van een bibliografische- en citatiedatabase valt uiteen in een aantal aspecten:

- het aantal ontsloten publicaties en documenten, gespecificeerd naar vakgebied
- dekkingsperiode voor seriële publicaties
- actualiteit van de dekking (hoe snel worden nieuwe publicaties opgenomen)
- aard van de data per document (titel, auteur, abstracts, keywords, verwijzingen, citaties etc.)

Informatie over dekking in dit onderzoek is afkomstig uit:

- Bronnen van de leverancier van Scopus zelf
- Vergelijking van gedekte tijdschriften op basis van ISSN
- Beoordeling ontbrekende tijdschrifttitels door UBU-vakspecialisten
- Vergelijking aantal citaties
- Interviews veelgebruikers onder de facultaire onderzoekers
- Webenquete

4.1 *Het aantal ontsloten documenten*

4.1.1 Totaal aantal documenten

Het totaal aantal documenten dat Scopus ontsluit bedroeg volgens eigen telling op 27 juni 2006 bijna 28 miljoen. Meer dan 95% daarvan bestaat uit tijdschriftartikelen. Boeken vormen nauwelijks een promille van het aantal records. Dit aantal records is kleiner dan de volledige versie van Web of Science (lopend van 1945-heden met 35 miljoen records (Jasco 2005)), maar beduidend groter dan de UBU-versie van WoS (lopend van 1988-heden, met 19 miljoen records (schatting op basis van Jasco 2005, p 1542)). Meer dan 50% van de documenten geïndexeerd in Scopus is gepubliceerd vóór 1996.

4.1.2 Aantal ontsloten tijdschriften

Scopus ontsluit met ruim 15.000 titels aanzienlijk meer tijdschriften dan Web of Science (bijna 9.000). Op zich is informatie over de miljoenen artikelen in die extra tijdschriften waardevol, maar het is natuurlijk wel de vraag wat de aard is van die extra tijdschriften. Lange tijd is het corpus tijdschriften ontsloten door WoS beschouwd als het topsegment van de tijdschriften. In dit rapport kijken we ook naar de kwaliteit van tijdschriften die Scopus wel ontsluit, maar WoS niet, en andersom. Verschillen hierin kunnen mede de keus voor een van beide bestanden bepalen.

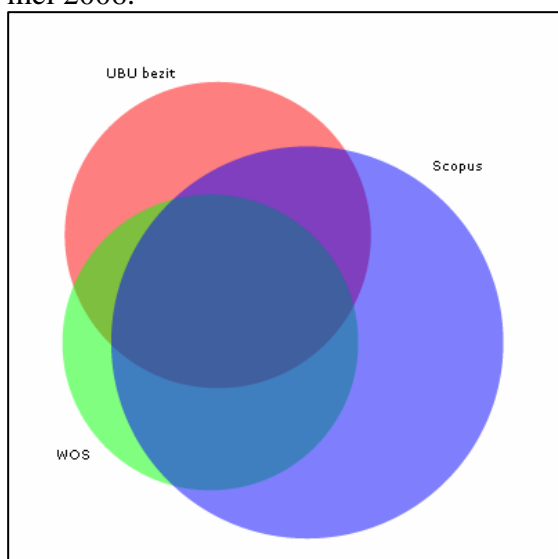
Het verschil tussen Scopus en WoS wat betreft aantal geïndexeerde tijdschriften is op zich interessant omdat Scopus duizenden titels meer dan WoS ontsluit op een wijze (met citatiegegevens) die alleen de WoS in die kwaliteit biedt. Andere bestanden met

citatiegegevens hebben ofwel veel minder professionele mogelijkheden (Google Scholar) of dekken slechts een of enkele vakgebieden (bv. Citeseer).

De grotere omvang in aantal gedekte tijdschriften bij Scopus is voor een belangrijk deel terug te voeren op het feit dat Scopus international gericht is. Het aandeel tijdschriften uit VS, Canada, VK, Nederland (Elsevier), Duitsland (Springer) en Zwitserland is in WoS 78% tegen 67% in Scopus (Pipp 2006). Scopus ontsluit 6 maal meer Chinese en driemaal meer Spaanse, Russische, Indiase, Poolse en Italiaanse tijdschriften dan WoS (Pipp 2006).

Het is interessant te zien hoe Scopus en WoS zich tot elkaar verhouden als ontsluitingsmiddel voor het bezit van de UBU. Helaas hebben we geen geïntegreerde tijdschriftenlijst van het gehele Utrechtse tijdschriftenbezit om te bezien welk deel daarvan in de databases voorkomt. Als we alleen kijken naar de tijdschriften die de UBU *digitaal* in haar bezit heeft (dat is ongeveer tweederde van het totale aantal lopende tijdschriften binnen UBU) zien we dat het verschil tussen de producten van ISI en Elsevier nog aanzienlijk is (figuur 4.1 en tabel 4.1). Met respectievelijk 53 en 64% ontsluit Scopus 11 procentpunten meer van de digitale UBU-titels. Zonder de EBSCO tijdschriften is dit verschil met 4 procentpunten aanzienlijk kleiner. Hier staat tegenover dat Scopus meer tijdschriften ontsluit die de UBU niet digitaal heeft dan WoS. Uit gesprekken blijkt dat onderzoekers dit vaak prettig vinden terwijl studenten het doorgaans als ballast zien.

Figuur 4.1 Overlap van digitale tijdschrifttitels UBU en titels in Web of Science en Scopus, mei 2006.



Bron: eigen onderzoek

Tabel 4.1. Overlap van digitale tijdschrifttitels UBU en titels in Web of Science en Scopus, mei 2006.

		Aantal tijdschriften	UBU	WOS	Scopus
inclusief Ebsco	UBU	9616		5142 (53%)	6162 (64%)
	WOS	8974	5142 (57%)		7505 (84%)
	Scopus	15785	6162 (39%)	7505 (48%)	
exclusief Ebsco		Aantal tijdschriften	UBU	WOS	Scopus
	UBU	7810		58%	62%
	WOS	8974	50%		81%
	Scopus	14191	34%	51%	

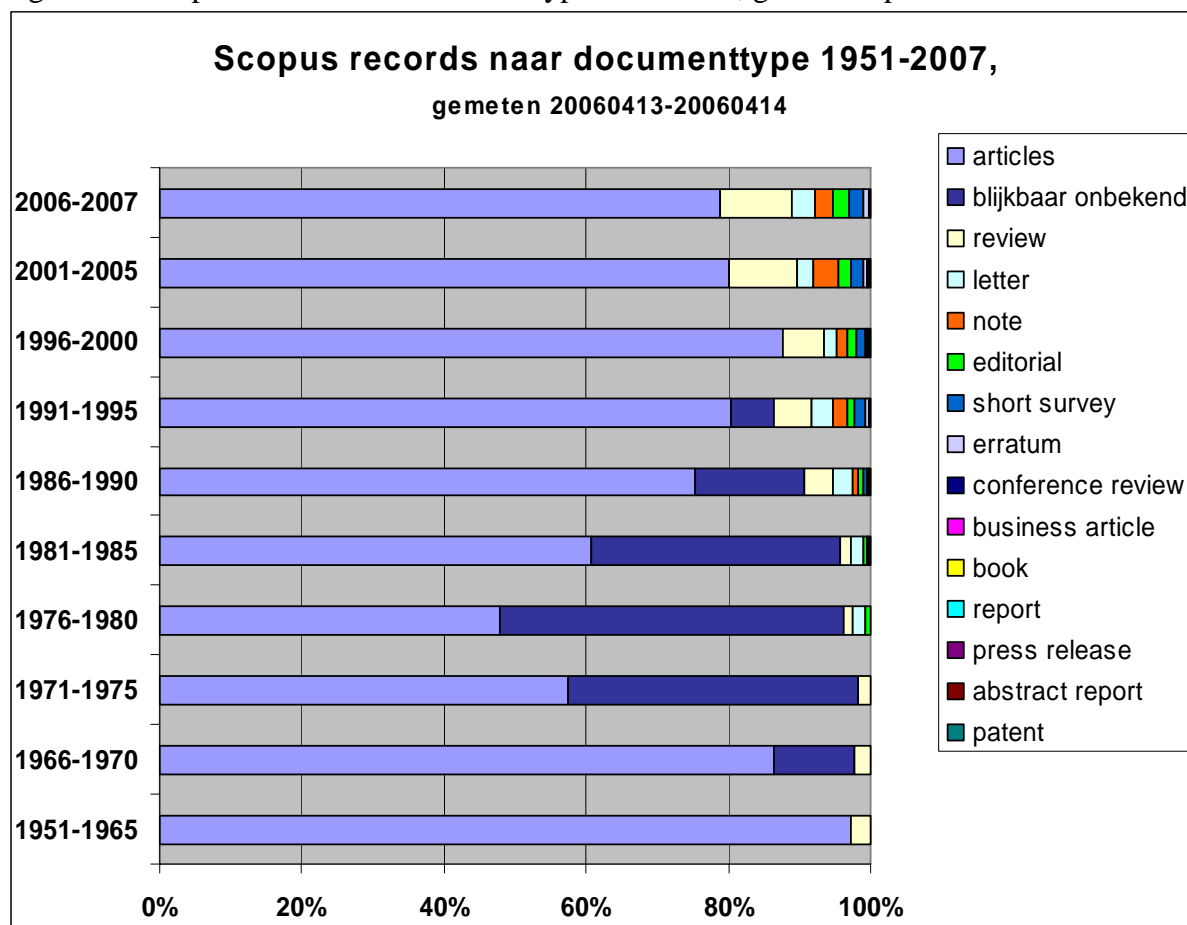
Bron: eigen onderzoek

Een belangrijke vraag is in hoeverre Scopus in staat is de opgenomen tijdschrifttitels ook daadwerkelijk compleet te maken. In de literatuur (Goodman & Deis 2005, Deis & Goodman 2006, Jascó 2006) wordt herhaaldelijk gewezen op het ontbreken van afleveringen en jaargangen. Hoewel dit ook in web of Science voorkomt, lijkt het bij Scopus veelvuldiger het geval. Dit behoeft nadere analyse en aandacht, niet in het minst van de makers van Scopus zelf.

4.1.3 Ontsloten documenten naar documenttype

Van de 28 miljoen records in Scopus is ruim 90% een beschrijving van een tijdschriftartikel. Dat betekent dus dat er toch nog enige miljoenen andersoortige bronnen beschreven worden. Vooral voor de recente periode zijn er vrij veel reviews, brieven, notities en surveys opgenomen (figuur 4.2). Veelal is dat materiaal dat wel uit tijdschriften afkomstig is. Het aantal niet-tijdschriftbronnen (boeken, rapporten, boekseries, congresbundels etc.) is met een kleine dertigduizend vrij klein. Op dat terrein bieden sommige vakindexen (Georef, PsycInfo) en Google Scholar veel meer. Overigens is de uitsplitsing naar publicatietype in de refine tools misleidend. Omdat voor oudere jaren bij een groot aantal records het publicatietype onbekend is (donkerblauw in figuur 4.2), wordt dat sterk onderschat. Het is dan ook af te raden bij zoekacties naar materiaal voor van 1996 te limiteren op artikelen als publicatietype, omdat daarmee ten onrecht miljoenen records worden buitengesloten. Dit is een belangrijke fout in de database die snel gerepareerd moet worden.

Figuur 4.2 Scopus records naar documenttype 1951-2007, gemeten april 2006



bron: eigen onderzoek

Vergelijking met Web of Science (tabel 4.2) laat zien dat Scopus voor een recent jaar duidelijk meer niet-artikel records bevat. Bij beide databases bestaat de bulk daarvan uit reviews. Ondanks deze voor Scopus voordelige totaalcijfers merkt Pipp (2006, 12) op dat dat voor specifieke tijdschriften, waaronder belangrijke in Scopus toch honderden niet-artikel bijdragen ontbreken. Controle bevestigt dit, en Scopus kan hier duidelijk nog verbeteringen doorvoeren.

Tabel 4.2. Documenttypen in WoS en Scopus, publicatiejaar 2005, gemeten op 20060414 met zoekactie op *relatively OR average*

	WoS	Scopus	WoS	Scopus
	absoluut		percentage	
alle typen	57018	74273	100,00%	100,00%
article	54650	66824	95,85%	89,97%
review	1989	6531	3,49%	8,79%
short suvey	0	356	0,00%	0,48%
note	0	277	0,00%	0,37%
business article	0	233	0,00%	0,31%
editorial	201	20	0,35%	0,03%
letter	58	13	0,10%	0,02%
conference review	0	12	0,00%	0,02%
erratum	0	6	0,00%	0,01%
abstract report	0	1	0,00%	0,00%
bibliographical item	1	-	0,00%	
book review	1	-	0,00%	
correction	24	-	0,04%	
meeting abstract	51	-	0,09%	
news item	14	-	0,02%	
reprint	24	-	0,04%	
software review	2	-	0,00%	

bron: eigen onderzoek

Een belangrijk gemis, zeker voor de sociale wetenschappen, sociale geografie & planologie en economie is het ontbreken van boekbesprekingen. Overigens is het wel goed dat boekseries in de bronnenlijst apart zichtbaar kunnen worden gemaakt. Scopus bevat bijvoorbeeld vrijwel alle afleveringen van de Nederlandse Geografische Studies - een reeks met vooral proefschriften - compleet met doorklikbare literatuurlijsten.

4.1.4 Ontsloten documenten en tijdschriften naar vakgebied

In geen multidisciplinair bestand is de dekking voor alle vakgebieden gelijk. Voor Scopus geldt dat humaniora bewust niet zijn opgenomen, omdat in deze vakgebieden tijdschriften minder belangrijk zijn. Alleen wetenschapsfilosofie is wel in Scopus opgenomen. Volgens Elsevier zelf heeft de nadruk bij de ontwikkeling in eerste instantie gelegen op STM (Science, Technology, Medicine) en daarnaast op Social Science (psychologie, sociologie, economie). Om te achterhalen in hoeverre claims worden waargemaakt hebben we tijdschriftenlijsten van verschillende bibliografische databases naast elkaar gelegd en op basis van ISSN gematched (tabel 4.3). Het ging hierbij puur om het voorkomen van overeenkomstige ISSN's, los van het aantal beschikbare jaren per tijdschrift. We vergelijken Scopus hier met één andere multidisciplinaire database, Web of Science en met 21 vakindexen (waarbij we EBSCO Academic Search Elite) voor het gemak bij de vakindexen hebben geschaard). De keuze voor

vakindexen is gemaakt op advies van vakspecialisten, maar beperkt door verkrijgbaarheid van volledige tijdschriftenlijsten met ISSN-aanduiding. Ook de Scopus-lijst bevatte helaas niet bij elke titel een ISSN waardoor de dekking van Scopus soms onderschat wordt.

Tabel 4.3. Dekking van titels uit vakspecifieke bibliografische databases in WoS en Scopus, met aanduiding van full text digitale beschikbaarheid ('UBU'), stand mei 2006.

Database	Aantal	% in UBU	% in WOS	% in SCOPUS	In beide
CSALISA	477	24,3%	15,7%	25,6%	14,7%
BIOSIS	3244	*	57,9%	82,8%	*
Georef	13345	7,2%	9,4%	13,2%	8,7%
Agricola	2215	28,2%	36,4%	40,5%	35,3%
CSAIPA	371	29,6%	34,5%	59,8%	34,0%
Philosophersindex	1194	25,6%	24,5%	15,7%	11,1%
ATLA	1549	24,2%	21,8%	10,1%	7,7%
Geobase	2065	43,3%	51,8%	87,4%	50,3%
EconLit	1077	41,1%	28,8%	45,4%	27,9%
SociologicalAbstracts	1777	42,7%	37,1%	48,7%	33,0%
PsycInfo	1990	52,8%	54,3%	70,5%	53,0%
CAB	7443	3,4%	38,0%	48,0%	37,2%
INSPEC	8966	18,7%	20,8%	29,3%	20,0%
MathSciNet	2697	22,0%	25,1%	30,1%	23,1%
LLBA	1555	31,9%	28,9%	26,4%	19,2%
Pubmed	7541	*	54,1%	60,3%	*
Embase	4865	47,0%	53,4%	91,6%	52,8%
SocIndex	3703	34,7%	29,6%	38,6%	26,6%
Eric	1264	46,7%	30,9%	40,0%	28,6%
EbscoAcademicSearch	7739	65,8%	56,4%	64,1%	50,2%
BHA	2244	7,8%	11,9%	2,5%	1,6%
MLA	4923	12,9%	14,0%	3,9%	3,0%

Bron: eigen onderzoek

* wordt nog berekend, beschikbaar in een nieuwe versie van dit rapport vanaf plm. 20060820

Voor 18 van de 22 lijsten van tijdschriften van vakindexen heeft Scopus een *hogere dekking* dan WoS. Opvallend zijn vooral de grote dekking bij farmacie (CASIPA), geowetenschappen (Geobase), economie (Econlit) en geneeskunde (EMBase). Voor geowetenschappen en geneeskunde is dit geen verassing: Geobase zowel als EMBase zijn Elsevier producten. Het roept andersom de vraag op waarom ten opzichte van deze vakindexen de dekking van Scopus niet de volledige 100% is. Volgens Elsevier zelf is de overlap van Scopus met Embase en Compendex wel 100%. Wij vermoeden dat het verschil verklaard wordt door ontbrekende ISSN's in de Scopus-lijst. Belangrijk is de gegevens in de tabel alleen te interpreteren in vergelijkende zin tussen WoS en Scopus. Vergelijking van de dekking van alleen Scopus of WoS met de verschillende vakindexen is lastig omdat deze laatste uit hun aard sterk verschillen in totaal aantal geïndexeerde seriële publicaties (tijdschriften). Voor Georef geldt bijvoorbeeld dat daar duizenden rapportenreeksen van geologische diensten tussen zitten. Een bestand als Econlit daarentegen richt zich sterk op louter reguliere tijdschriften (en boeken), waardoor de vergelijking met multidisciplinaire bestanden tot hogere percentages leidt. Tenslotte is *lagere dekking* van Scopus vergeleken met Web of Science op het terrein van de humaniora (en de restrubriek 'algemeen') evident. Deze verschillen tussen Scopus en WoS naar vakgebied komen overeen met onze bevindingen op basis van citaties in beide databases (paragraaf 4.6).

Vergelijking met Google Scholar is hier moeilijk, omdat dat Google voor bestand geen lijst van brontijdschriften publiceert. Uit onderzoek is wel gebleken dat Google Scholar de beste dekking heeft voor tijdschriften uit de medische en biomedische wetenschappen, sterk wisselend voor de non-life natuurwetenschappen (goed voor informatica en chemie, minder voor wiskunde en aardwetenschappen), gemiddeld voor de sociale wetenschappen en economie en vrij slecht voor de humaniora (Neuhaus 2006). Daarbij is Google Scholar overigens sterker in ontsluiting van Open Access dan non-Open Access, relatief sterker in Engels dan andere talen, en sterker in de ontsluiting van tijdschriften bij multidisciplinaire grote uitgeversplatforms dan in tijdschriften in bibliografische databases.

Ook hebben we een vergelijking gemaakt van de dekking van tijdschriften die binnen de UU full text digitaal beschikbaar zijn in Scopus, WoS en de eigen Utrechtse zoekmachine van Omega (tabel 4.4). Hierbij maken we gebruik van de onderwerpsindeling van tijdschriften in Omega, die helaas nog wel wat te wensen overlaat. Ook zijn hier per onderwerp alleen die tijdschriften meegenomen die er primair aan gekoppeld zijn. Dat wil zeggen dat een tijdschrift dat in twee rubrieken is geplaatst, alleen voor het eerste meetelt. Belangrijk is te beseffen dat in de onderliggende gegevens niet altijd duidelijk is of een tijdschrift volledig of selectief wordt ontsloten in de databases en tevens dat hier het aantal jaargangen dat per tijdschrift is opgenomen geen rol speelt.

Tabel 4.4. Dekking van full text in UBU beschikbare tijdschrifttitels in WoS, Scopus en de eigen Omega zoekmachine, stand mei 2006, hoogste waarde gemarkeerd..

	UBU	in WOS		in Scopus		in Omega ZM	
Aardwetenschappen	299	222	74%	242	81%	195	65%
Algemeen	298	90	31%	103	35%	179	60%
Biologie	1043	768	74%	864	83%	612	59%
Diergeneeskunde	114	88	77%	93	82%	42	37%
Economie	659	263	40%	402	61%	429	65%
Farmacie	125	85	68%	101	81%	74	59%
Geneeskunde	2137	1234	58%	1784	83%	1263	59%
Godgeleerdheid	144	47	33%	12	8%	71	49%
Landbouwwetenschappen	82	56	68%	67	82%	53	65%
Letteren	952	430	45%	171	18%	418	44%
Milieukunde	129	62	48%	90	70%	88	68%
Natuur- en Sterrenkunde	464	315	68%	349	75%	364	78%
Rechtsgedeerdheid	217	47	22%	75	35%	104	48%
Soc. Geografie & Planol.	156	61	39%	100	64%	88	56%
Scheikunde	346	246	71%	261	75%	218	63%
Sociale Wetenschappen	1231	506	41%	666	54%	805	65%
Techniek	295	169	57%	218	74%	215	73%
Wijsbegeerte	93	51	55%	17	18%	39	42%
Wiskunde en Informatica	832	402	48%	496	60%	432	52%
Totaal	9616	5142	53%	6111	64%	5689	59%

Bron: eigen onderzoek

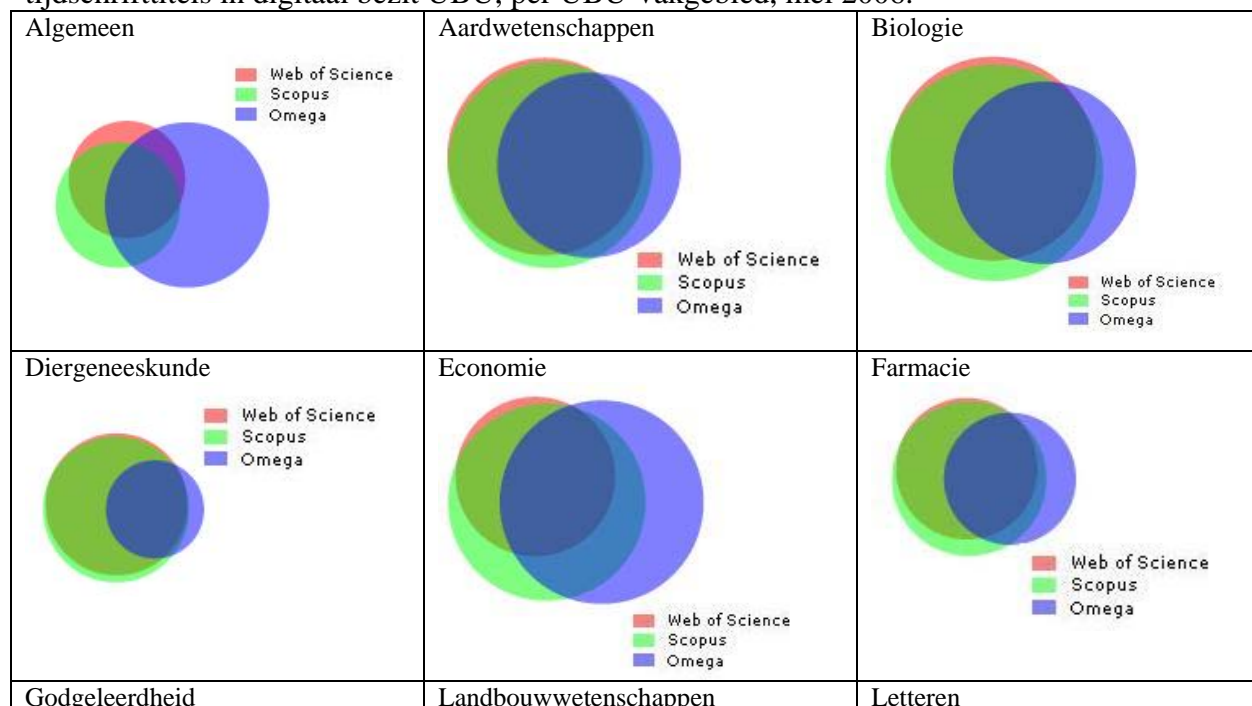
Met de hier gehanteerde beperking tot die tijdschriften die de UBU full text kan aanbieden is de grotere dekking van Scopus duidelijk: overall dekt Scopus met 64% ruim tien procentpunten meer van ons digitale tijdschriftenbezit. Met uitzondering van letteren, wijsbegeerte en godgeleerdheid (volgens de Omega indeling) laat Scopus ook hier in alle vakgebieden een ruimere dekking zien dan WoS. In 5 UBU-vakgebieden (waaronder ook 'algemeen') is de dekking van de Omega zoekmachine het grootste van deze drie

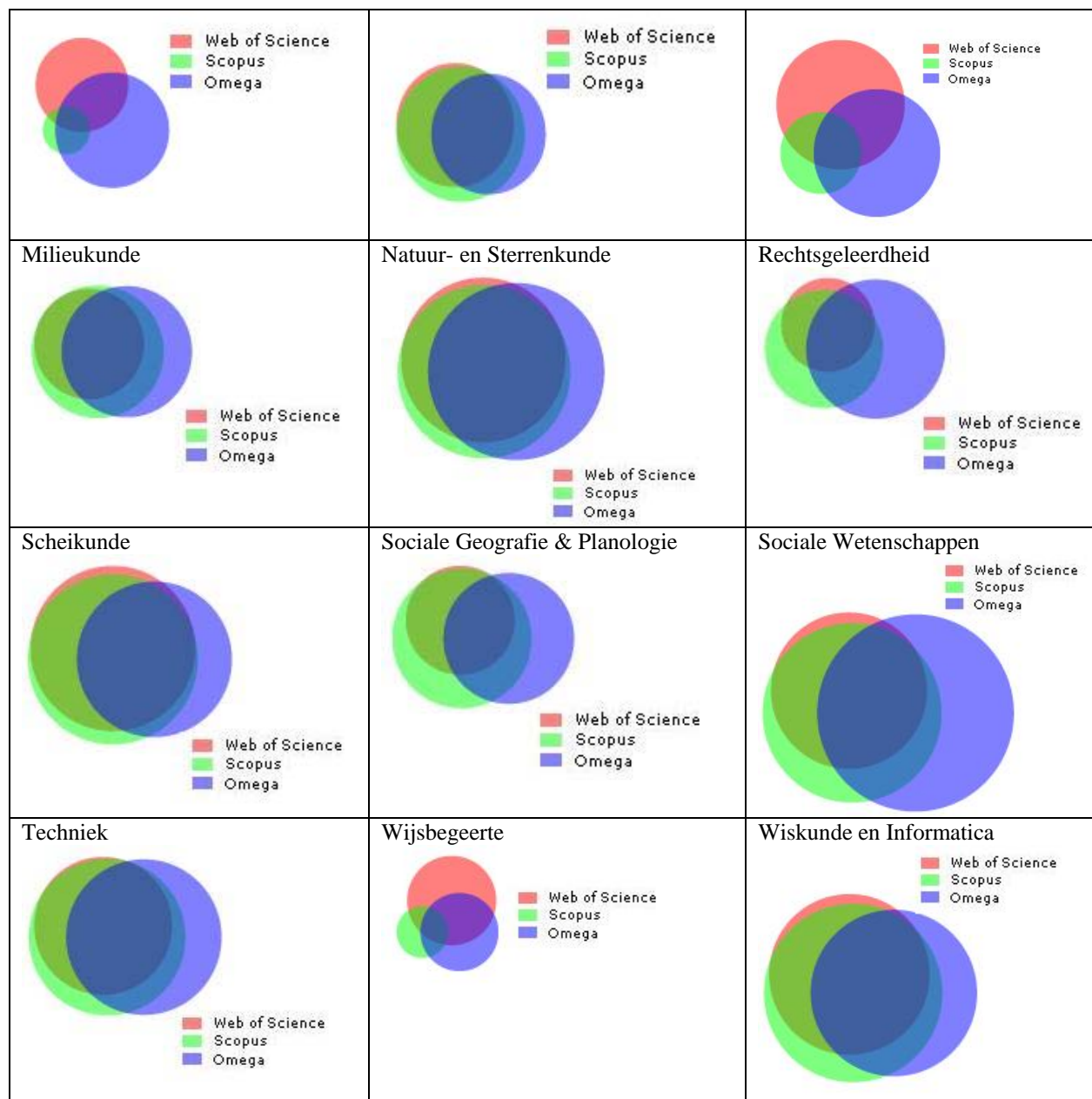
multidisciplinaire ingangen. Overigens is het belangrijk aan te merken dat bij de full text tijdschriften van de UBU alle tijdschriften zijn meegenomen, inclusief die met een moving wall voor recentere jaren (JSTOR, PCI) en inclusief vrij grote aantallen niet wetenschappelijke tijdschriften uit EBSCO ASE. Die laatste vormen vooral bij sociale wetenschappen, techniek en economie een relatief groot aandeel. Dat maakt dat daar de eigen Omega zoekmachine goed uit de verf komt, want daarin zijn alle full text titels uit EBSCO ASE ontsloten. De eigen Omega zoekmachine is ook sterk in vakgebieden waar een groot deel van de tijdschriften zich concentreert bij enkele van de grootste uitgevers (Elsevier, Springer en Wiley), waarvan ook alles in Omega ontsloten wordt. Dit geldt bijvoorbeeld voor milieukunde.

Voor een goede beoordeling van de dekking van tijdschriften van specifieke vakgebieden is het van belang om naast de aantallen tijdschriften ook te kijken naar de overlap van gedekte tijdschriften. Als twee databases op een vakgebied evenveel titels indexeren is de mate van overlap bepalend voor de keus om één van beide of beide databases te kiezen. Indexeren databases voor een vakgebied niet evenveel titels, maar is er wel volledige overlap, dan ligt het voor de hand te kiezen voor de database met het grootste aantal titels. Bij de meeste vakgebieden is er echter sprake van én onvolledige overlap én verschil in aantal gedekte titels (figuur 4.3).

Overall genomen is duidelijk dat Scopus qua ontsluiting van ons bezit veel meer waarde toevoegt aan de Omega-zoekmachine dan WoS. Met uitzondering van de humaniora is het aantal titels dat van de drie databases alleen in Scopus voorkomt in alle vakgebieden groter dan het aantal dat voor WoS uniek is. De overlap met de Omega-zoekmachine is meestal veel kleiner dan die van Scopus en WoS onderling. De kleinste onderlinge overlap WoS-Scopus zien we bij de humaniora en de rubriek algemeen en daarnaast bij rechtsgeleerdheid, bij economie en sociale geografie & planologie en bij de sociale wetenschappen en wiskunde & informatica. Bij alle overige vakken, allemaal ‘physical’ of ‘life’ sciences, is er vrij sterke overlap. Belangrijk is om voor elk vakgebied te bepalen hoe belangrijk de titels in het niet onderling overlappende gebied van WoS en Scopus zijn.

Figuur 4.3. Overlap van indexering door WoS, Scopus en de Omega-zoekmachine van tijdschrifttitels in digitaal bezit UBU, per UBU-vakgebied, mei 2006.





bron: eigen onderzoek

De exacte aantallen tijdschriften van het full-text UBU-bezit die alleen door Scopus, alleen door WoS of door geen van beide zijn opgenomen verschillen dus sterk per vakgebied (tabel 4.5).

Tabel 4.5 Tijdschrifttitels uit het digitale bezit UBU die niet in Scopus, WoS of niet in beide voorkomen, per UBU-vakgebied, juni 2006.

vakgebied	niet in WoS, wel in Scopus	niet in Scopus, wel in WoS	in Scopus noch WoS
Aardwetenschappen	20	5	57
Algemeen	14	21	191
Biologie	63	29	212
Diergeneeskunde	4	2	22
Economie	119	12	227
Farmacie	14	2	26
Geneeskunde	416	21	487

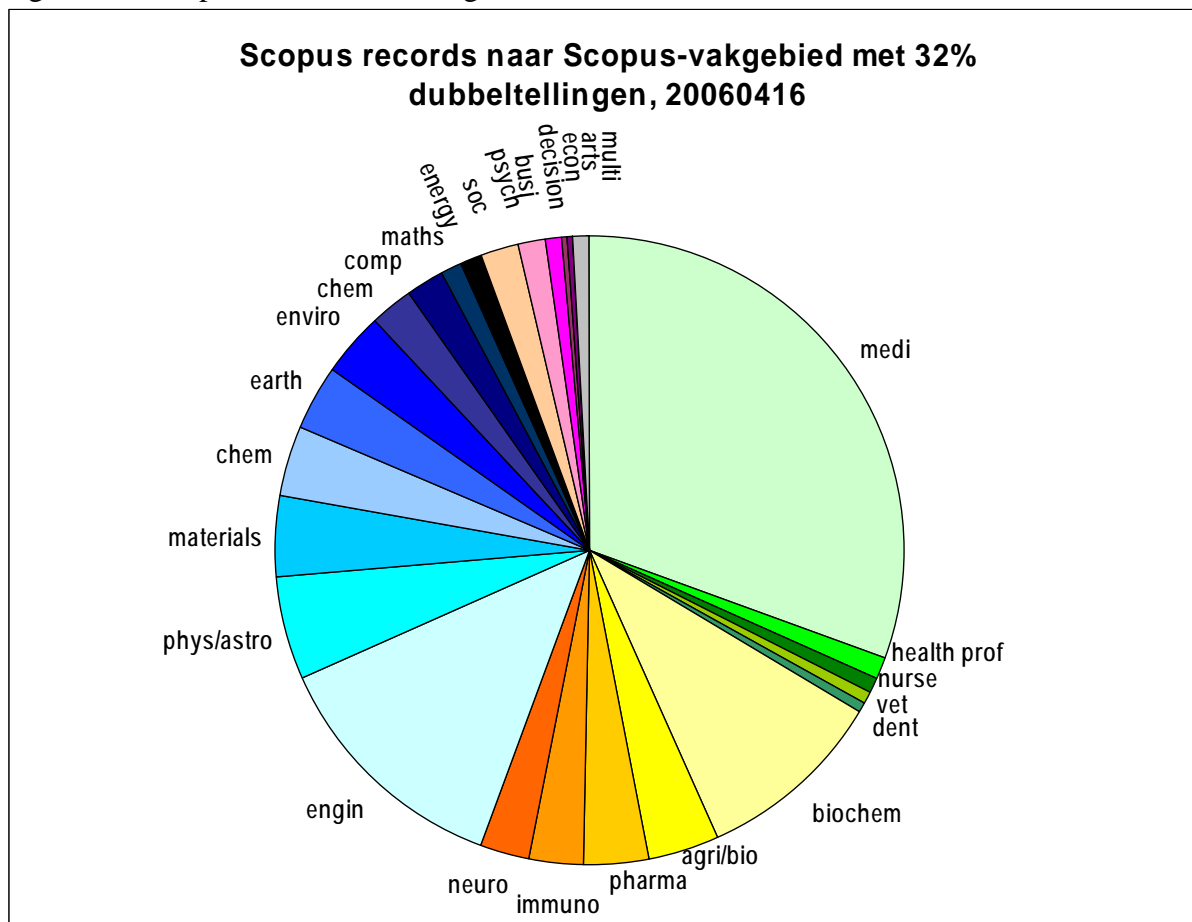
Godgeleerdheid	4	42	93
Landbouwwetenschap	8	2	18
Letteren	36	229	486
Milieukunde	21	0	46
Natuur- en sterrenkunde	19	14	129
Rechtswetenschap	13	6	157
Scheikunde	18	14	82
Sociale geografie & Planologie	30	1	65
Sociale Wetenschappen	104	33	621
Techniek	31	2	95
Wijsbegeerte	3	43	38
Wiskunde & Informatica	66	24	364

bron: eigen onderzoek

Vakspecialisten uit een aantal vakgebieden hebben gekeken naar deze lijsten met niet-overlappende titels van WoS en Scopus en beoordeeld in welke mate daar cruciale titels tussen zitten (voor door vakspecialisten als zodanig gemarkeerde titels zie bijlage I). De UBU-vakspecialisten hebben cruciale titels gemarkeerd niet om in kwantitatieve zin te oordelen over Scopus en WoS, maar om over ontbrekende titels te communiceren met de achterban in de faculteiten en met de leveranciers van de bestanden (Elsevier en Thomson-ISI).

Tenslotte kijken we naar het totaal aantal records per vakgebied (figuur 4.4). Hierbij is wel sprake van dubbel tellingen, aangezien een tijdschrift door Scopus soms onder meer dan één vakgebied wordt geschaard.

Figuur 4.4. Scopus records naar vakgebied



bron: eigen onderzoek

Sterker dan bij het aantal tijdschriften blijkt hier de sterkere STM-gerichtheid van de Scopus database. Die sterke nadruk op STM vergeleken met een meting gebaseerd op aantallen tijdschriften heeft te maken met het gemiddeld veel grotere aantal artikelen per jaargang in STM-tijdschriften en met het feit dat juist voor deze vakken de dekking van Scopus verder teruggaat en er dus meer jaargangen geïndexeerd zijn.

4.2 Dekking op basis van voorbeeldzoekacties

Naast tellingen van geïndexeerde tijdschriften geven concrete zoekacties een goede indruk van de omvang van de twee citatiedatabases. Het kan een ander beeld opleveren doordat dan aantallen *records* worden geteld. Daarmee speelt het aantal per tijdschrift geïndexeerde artikelen een rol, dat afhankelijk is van het aantal gepubliceerde artikelen in die tijdschriften en het aantal gedekte jaren in de database. De resultaten (tabel 4.6) leveren inderdaad een van de eerdere indrukken afwijkend en divers beeld op.

Tabel 4.6 Zoekresultaten van drie zoekacties in default velden, per vakgebied, aantallen records in Scopus als % van WoS, april 2006

Vakgebied	exacte zoekstring	1988-1995	1996-2005	1988-2005
aardwetenschappen	geophysic* AND geolog*	342	264	287
aardwetenschappen	groundwater AND monitoring AND model*	277	193	206
aardwetenschappen	foraminiferal AND "north Sea"	107	80	88
Biologie	(plant* OR animal* OR organism*) AND genera	186	295	269
Biologie	learn* and songbirds*	111	135	132
Biologie	root pattern OR "root patterns" OR "root patterning"	215	322	294
diergeneeskunde	veterinary	95	144	130
diergeneeskunde	embryogenesis AND bovine	61	44	50
diergeneeskunde	"animal diseases" AND vaccination	383	283	303
economie	"foreign direct investment" AND competit*	123	131	130
economie	("early modern" OR "post-war") AND econom*	188	120	135
economie	firm* AND merger* AND market*	38	170	144
farmacie	pharmac* AND receptor*	152	95	109
farmacie	polymers AND (liposomes OR "drug delivery systems")	264	223	230
farmacie	"drug targeting" AND "controlled release"	150	164	161
geneeskunde	cancer AND neuro*	162	96	108
geneeskunde	"lymphocyte development" AND thymus	151	63	79
geneeskunde	"endoplasmic reticulum" AND hormone*	110	74	87
milieukunde	enviro* AND pollut*	468	300	334
milieukunde	"food webs"	64	69	68
milieukunde	innovat* AND energ* AND biomass*	170	298	274
N&S:natuurkunde	"string theory"	35	61	55
N&S: natuurkunde	"condensed matter" AND optic*	102	111	110
N&S: natuurkunde	stratocumulus AND "boundary layer"	59	53	54
N&S: sterrenkunde	telescop* OR asteroid* OR supernova* OR interstellar	59	100	89
N&S: sterrenkunde	magnetohydrodynamic* AND plasma*	105	124	120
N&S: sterrenkunde	"stellar winds" AND nebulae	41	91	73

scheikunde	molecular AND aromatic	121	149	142
scheikunde	"protein folding" AND (molecular chaperones OR Hsp90 chaperone)	164	184	181
scheikunde	("phase behaviour" OR "phase behavior") AND (colloids OR rods)	205	109	121
sociale geografie en planol.	(geographical OR spatial) AND (urban OR economic)	324	178	204
sociale geografie en planol.	regional AND evolutionary* AND (business* OR compan* OR econom*)	156	115	119
sociale geografie en planol.	"southern africa" AND develop* AND econom*	383	329	342
SW: antropologie	anthropo*	119	129	127
SW: antropologie	(trauma* OR violen*) AND (ethnic* OR ethno* OR societ*)	138	131	132
SW: antropologie	migrat* AND ethnic*	376	153	194
SW: psychologie	psychol*	152	157	155
SW: psychologie	(aggression OR criminality) AND psycho*	162	110	120
SW: psychologie	neuropsycholog* AND psychopatholog* AND cogniti*	110	86	90
SW: sociologie	sociolog*	58	92	80
SW: sociologie	gender* AND household* AND (labor OR labour OR work*)	118	96	100
SW: sociologie	"life course" OR "life courses"	49	77	72
W&I: informatica	"computational complexity" AND Bayesian	480	488	487
W&I: informatica	computational AND geometr* AND virtual	208	622	587
W&I: informatica	programming AND distributed	613	428	475
W&I: wiskunde	(algebra* OR arithmetic*) AND calculus	87	137	123
W&I: wiskunde	"Lie algebras"	8	68	49
W&I: wiskunde	ocean AND (eigenfunction* OR eigenvector*)	169	210	197

bron: eigen onderzoek; NB. groen=Scopus meer dan 10% meer resultaat dan WoS, rood=WoS meer dan 10% meer resultaat dan Scopus, geel=verschil tussen aantal zoekresultaten Scopus en WoS 10 procent of minder.

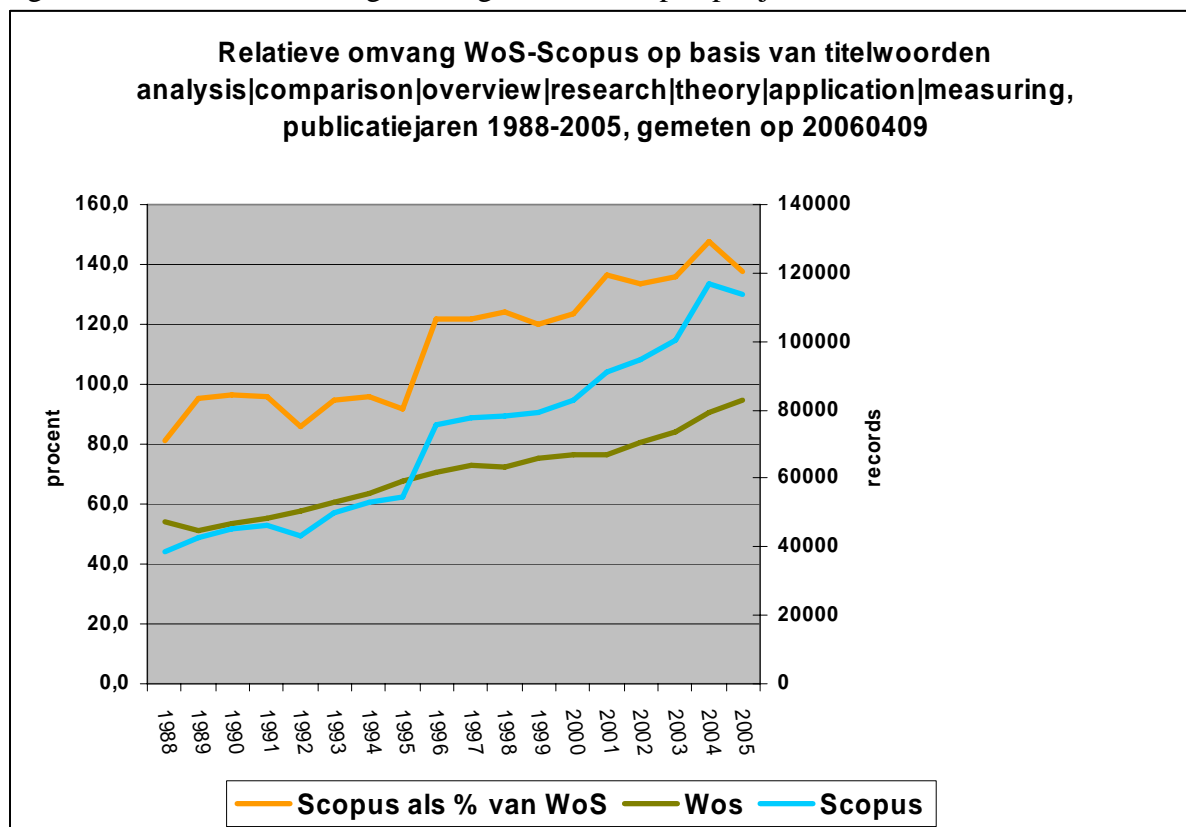
Bij de resultaten van de zoekacties vallen een paar dingen op. In het algemeen geeft Scopus duidelijk bij de meerderheid van de zoekacties een groter resultaat dan Web of Science. Maar dit beeld is niet compleet zonder te wijzen op drie andere opvallende zaken. Ten eerste het niet eenduidige beeld bij diergeneeskunde en geneeskunde. Op sommige zoekacties scoort hier WoS beter, op andere Scopus. Ten tweede de minder goede score van Scopus bij sociologie, natuurkunde en sterrenkunde. Ten derde de slecht verklaarbare extreme cijfers van enkele zoekacties bij wiskunde en milieukunde.

Uiteraard geldt bij het kleine aantal zoekacties per vakgebied een caveat voor al te snelle conclusies. Ook de automatische stemming van Scopus (waardoor bijvoorbeeld meervoudsvormen worden meegenomen) en de Keywords-Plus van WoS hebben een effect op de aantallen dat lastig te verdisconteren is bij de interpretatie.

4.3 Dekkingsperiode

De totale omvang van Scopus in vergelijking tot die van WoS gemeten door de tijd is alleen indirect te bepalen. Voor WoS zijn namelijk geen totalen per jaar bekend. Ook zoekacties op alleen publicatiejaar laat WoS niet toe. Als proxy hebben we gewerkt met een zoekactie op algemene, niet vakspecifieke titelwoorden waarmee we ook in recente jaren geen resultaat krijgen dat in WoS meer dan 100,000 treffers oplevert, aangezien WoS dan het exacte aantal niet weergeeft. Het resultaat van deze zoekactie staat in figuur 4.5. Vóór 1996 is Scopus volgens deze schatting 5-15% kleiner dan WoS, na 1996 20-45% groter.

Figuur 4.5. Relatieve omvang dekking WoS en Scopus per jaar, 1988-2005



bron: eigen onderzoek

Het feit dat Scopus zeker verder teruggaat dan 1996 blijkt duidelijk ook uit figuur 4.6. Ruim 50% van de documenten is vóór 1996 gepubliceerd. Wel is het zo dat alleen voor de documenten vanaf publicatiejaar 1996 citatiegegevens getoond worden. Oudere documenten fungeren wel ook als brondocument voor citatietellingen.

Tabel 4.7 Elsevier databases waarop Scopus voor een aanzienlijk deel is gebaseerd.

database	Dekkingsperiode
Medline (via EMBase)	1966-...
EMbase	1970-...
Compendex	1970-...
World Textile Index	1970-...
Fluidex	1974-...
Geobase	1980-...
Biobase	1994-...

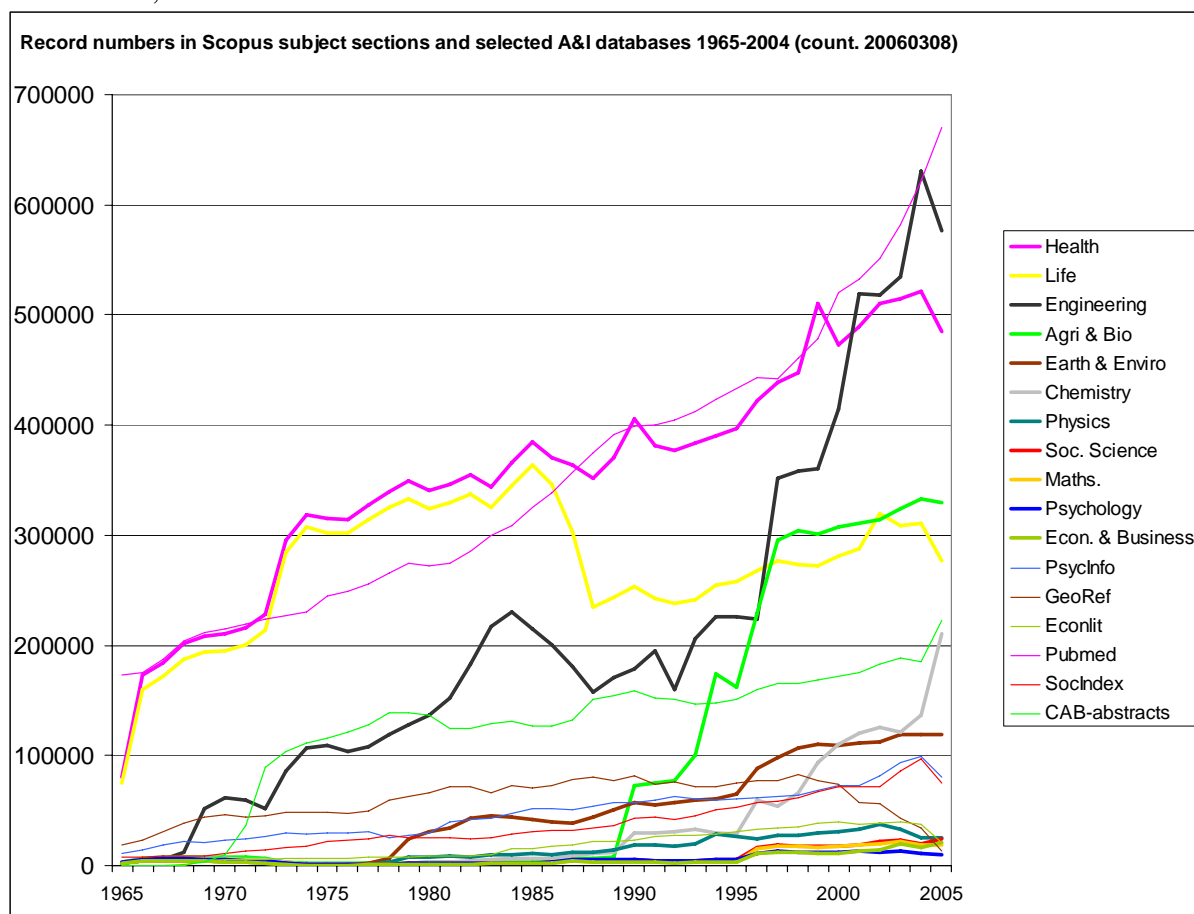
bron: Goodman & Deis 2005

Dekking vóór 1996 lijkt vooral terug te voeren op de databases die Elsevier reeds had: EMBase, Biobase, Geobase en Compendex, naast kleinere databases op het gebied van vloeistoffen (Fluidex), Oceanologie (Ocean Base) en textiel (World Textiles) (tabel 4.6). Daarmee is de dekking voor life, health, agri/bio en earth/enviro en ook techniek (engineering) verder terug in de tijd goed. De dekking voor psychologie, economie en maatschappijwetenschappen, maar ook voor wiskunde daarentegen, is vóór publicatiejaar 1996 relatief zeer beperkt. De inhoud van Scopus vóór 1980 is eigenlijk vooral nog biomedisch.

Vóór 1966 is de dekking in alle vakgebieden minimaal. Dit is een verschil met WoS, dat in onze versie weliswaar niet verder teruggaat dan 1988, maar in principe (voor het science deel) teruggaat tot 1945 en sinds kort zelfs voor ruim 200 tijdschriften tot 1900.

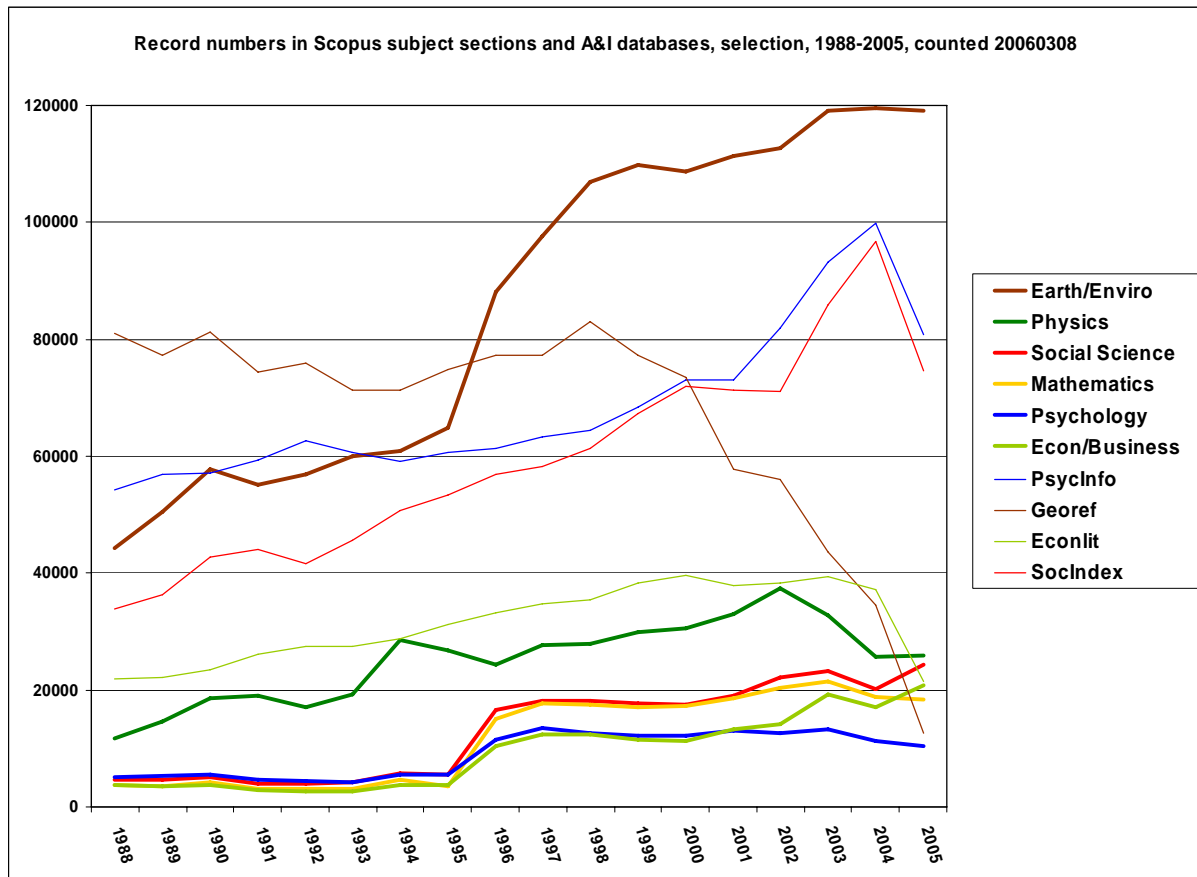
In Utrecht hebben we wel op andere manieren digitale ontsluiting van oude tijdschriften in JSTOR, PCI, bij de uitgeversplatforms Science Direct en SpringerLink, in sommige vakindexen (bijvoorbeeld Georef, PsycInfo en Zentralblatt MATH), voor sommige tijdschriften in Online Contents en ook voor materiaal van sommige uitgevers via Omega. De beperkte dekking voor met name de sociologie in de periode voor 1996 (goed zichtbaar in de uitvergroting van de periode 1995-2005 in figuur 4.7) tenslotte, is een minpunt in vergelijking met de WoS. Daarbij moet wel worden aangetekend dat deze cijfers zijn van maart 2006. In de periode daarna is Scopus selectief doorgedaan met de zogenaamde 'backfill'.

Figuur 4.6. Aantallen records van de verschillende Scopus-vakgebieden en van geselecteerde vakindexen, 1965-2005.



bron: eigen onderzoek; NB. Sinds het uitvoeren van deze telling is de vakgebiedsindeling van Scopus verfijnd en gewijzigd, waardoor deze tellingen niet meer reproduceerbaar zijn.

Figuur 4.7. Aantallen records van de verschillende Scopus-vakgebieden en van geselecteerde vakindexen, 1965-2005 (uitvergroting van deel van figuur 4.6).



bron: eigen onderzoek

Een indicatie van de 'backfill', het opnemen van publicaties van vóór 1996 is de verhouding tussen het aantal recente en het aantal oudere records (tabel 4.8).

Tabel 4.8. Verhouding tussen records 1986-1995 en 1996-2005 als indicatie voor 'backfill' van de database, gemeten 20060627, op basis van de nieuwe gedetailleerde Scopus vakgebiedsindeling (met dubbelstellingen).

Scopus vakgebied	records 1996-2005	records 1986-1995	1986-1995 als % van 1996-2005
Medicine	3871121	3002118	78
Environmental Science	514086	385134	75
Pharmacology, Toxicology and Pharmaceutics	503491	370771	74
Dentistry	66562	48766	73
Immunology and Microbiology	477611	322546	68
Biochemistry, Genetics and Molecular Biology	1565869	1039927	66
Earth and Planetary Sciences	576363	381365	66
Neuroscience	396825	261343	66
Nursing	125955	78039	62
Health Professions	199752	121077	61
Engineering	1964645	1160747	59
Multidisciplinary	123439	59812	48
Veterinary	131972	57236	43

Psychology	257340	105019	41
Energy	233501	90820	39
Materials Science	921774	328310	36
Agricultural and Biological Sciences	958706	278408	29
Decision Sciences	59071	16850	29
Chemical Engineering	552390	157464	29
Computer Science	509072	145008	28
Physics and Astronomy	1379044	341580	25
Chemistry	935545	222622	24
Social Sciences	492138	114669	23
Arts and Humanities	44248	7460	17
Economics, Econometrics and Finance	115481	14059	12
Business, Management and Accounting	230888	25477	11
Mathematics	405098	39124	10

bron: eigen onderzoek

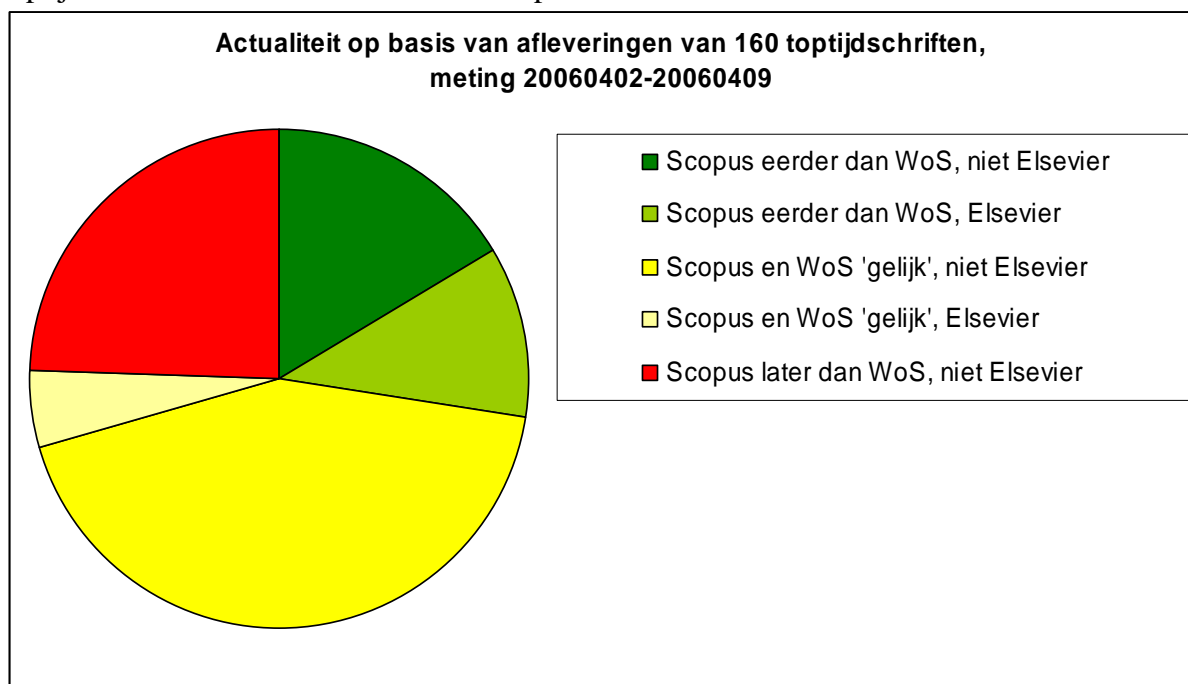
Het is duidelijk dat vakgebieden waarvoor er geen onderliggende database is relatief weinig oudere records in Scopus hebben: sociologie, maatschappijwetenschappen, maar ook scheikunde, wiskunde en natuur- en sterrenkunde. Psychologie heeft ondanks het ontbreken van een onderliggende Elsevier-database op dat terrein een redelijke backfill, waarschijnlijk gebaseerd op tijdschriften die (ook) in EMBASE zitten.

4.4 Actualiteit

Het vermeende (Goodman & Deis 2005) verschil in actualiteit van de databases Scopus en WoS in het nadeel van Scopus, blijkt thans niet uit onze analyse van geïndexeerde afleveringen van 160 toptijdschriften. Voor deze analyse hebben we de twee titels met de hoogste impact factor uit 80 voor onze universiteit belangrijke rubrieken van de Journal Citation Reports genomen (55 science en 25 social science). Het is duidelijk dat voor bijna de helft van deze tijdschriften Scopus en WoS even ver zijn; voor iets meer dan een kwart is Scopus verder, voor iets minder dan een kwart is WoS verder (figuur 4.8). Te verwachten viel dat bij de tijdschriften waarvoor Scopus verder is dan WoS een disproportioneel aantal Elsevier-titels zit.

Pipp (2006, p.14) heeft voor een kleiner aantal tijdschriften een vergelijkbare test uitgevoerd, maar daarbij tevens gekeken hoe groot de achterstand van zowel WoS als Scopus is op de publicatie van het tijdschrift op het uitgeversplatform (Blackwell Synergy, SpringerLink, Scienc Direct, etc.). Zij komt tot de conclusie dat Scopus iets actueler is dan Web of Science, maar dat Scopus voor een klein aantal tijdschriften ver achterloopt. Blijkbaar heeft Scopus de workflow nog niet voor alle titels op orde.

Figuur 4.8. Actualiteit van Scopus en WoS op basis van beschikbaarheid van afleveringen van toptijdschriften uit de Journal Citation Reports.

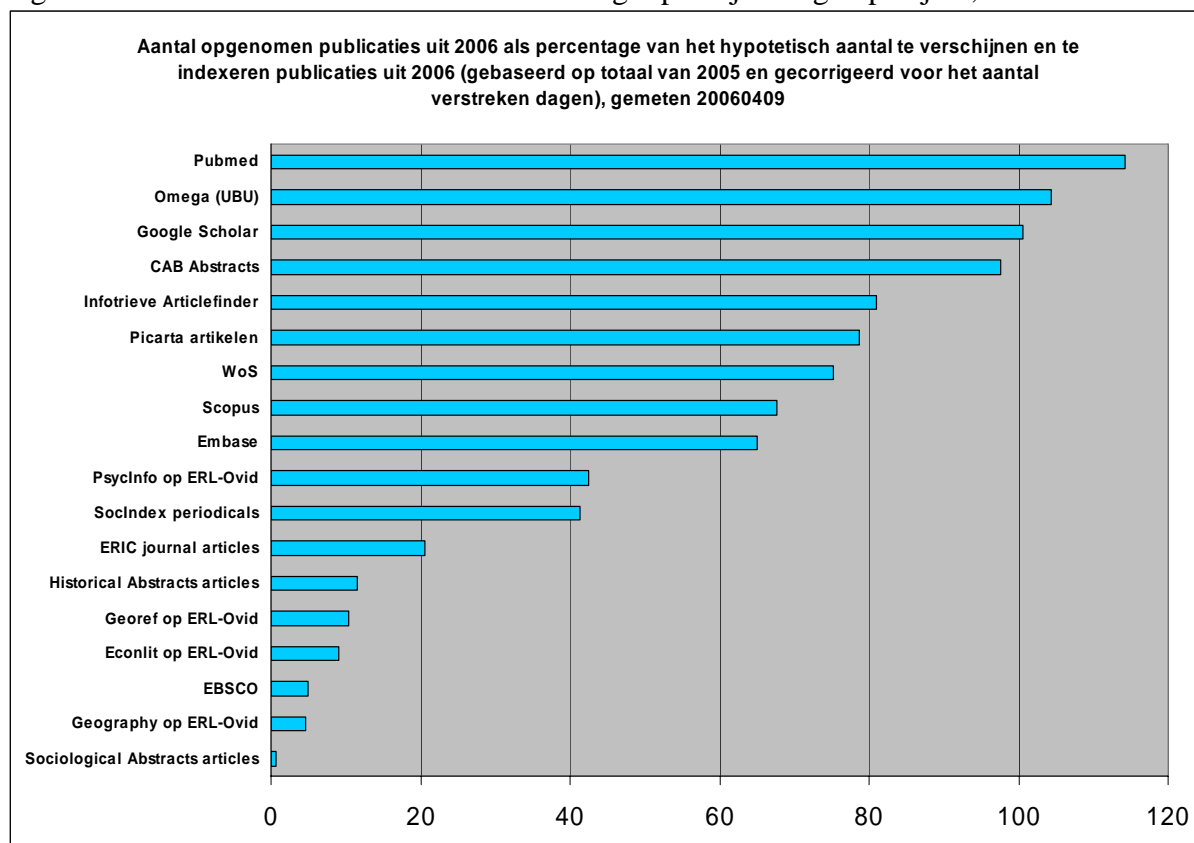


bron: eigen onderzoek

Een andere manier om naar actualiteit te kijken is het aantal records gepubliceerd in het lopende jaar te vergelijken met dat gepubliceerd in het afgelopen jaar, corrigerend voor het deel van het lopend jaar dat nog moet komen. Gezien de tijd die nodig is voor verwerking is het te verwachten dat na bijvoorbeeld 4 maanden zelden een derde van het te verwachten aantal titels van dat jaar dat opgenomen zal zijn in de database. Daarnaast is er in omgekeerde richting juist een relatief groot aantal geïndexeerde documenten te verwachten, door de absolute groei van het aantal publicaties van jaar tot jaar. Het beeld in figuur 4.9 moet dan ook alleen in vergelijkende zin worden geïnterpreteerd. Een niveau van 100 geeft aan dat als het aantal publicaties in 2006 gelijk zou zijn aan dat in 2005 de makers van de database hypothetisch precies op schema zitten om aan het eind van het lopende jaar alle publicaties van dit jaar opgenomen te hebben. Het gaat om het verschil van de mate waarin databases afwijken van dat niveau van 100. Verklaring van de verschillen zit in ofwel daadwerkelijk tragere verwerking ofwel een groter aandeel traag verschijnende titels in het bestand. Alleen als de databases verschillen in de mate waarin ze vanaf 2006 nieuwe titels hebben opgenomen waarvan ze ook alleen 2006 indexeren, treedt er werkelijke vertekening op. Het zelfde geldt voor de situatie waarin databases sterk zouden verschillen in de mate waarin per 2006 gestopt is met het indexeren van bepaalde titels.

Van de grote databases zit Scopus duidelijk in de middenmoot, wat lager dan WoS. We zagen al dat afgaande op geïndexeerde afleveringen van toptijdschriften geen groot verschil in actualiteit is. Het is lastig hier te concluderen of er ook sprake is van daadwerkelijk tragere verwerking. Voor een deel zal het lagere niveau van Scopus namelijk ook te maken hebben met het feit dat Scopus vanuit z'n onderliggende databases voor sommige vakgebieden ook veel minder snel publicerende tijdschriften indexeert. Omgekeerd stelt de WoS een strak publicatieregime juist als voorwaarde voor opname van tijdschriften. Ook bij de echte vakindexen kunnen we de invloed van traag verschijnende tijdschriften verwachten, naast het feit dat sommige van deze bibliografieën ook minder vaak geactualiseerd worden.

Figuur 4.9 Actualiteit van databases: verhouding lopend jaar/afgelopen jaar, 2006.



bron: eigen onderzoek

4.5 Aard van de opgenomen gegevens per document

Naast dekking en actualiteit is ook de informatie opgenomen per verwijzing een belangrijk punt in de vergelijking van bibliografische databases. Voor professioneel gebruik zit daarin vaak belangrijke meerwaarde ten opzichte van simpeler zoekingen als Google Scholar of Online Contents, zoals adresgegevens van auteurs en (doorzoekbare) keywords. Vergelijking van één record in verschillende databases (tabel 4.9) levert een goede eerste indruk op van de verschillen.

Uit vergelijking van de velden waarop de ingangen verschillen blijkt duidelijk dat de grote citatiedatabases het meest compleet zijn wat betreft hoeveelheid informatie per record. In Scopus zowel als in WoS zijn de records uitgebreider dan in de meeste andere bibliografische databases. Veelal geldt dat ook voor het aantal keywords. Toch zijn er onderling grote verschillen in de beschikbaarheid van keywords in Scopus en WoS.

Scopus heeft naast de auteurstrefwoorden vaak trefwoorden uit gecontroleerd vocabulaire afkomstig uit onderliggende databases als Compindex, Geobase en EMBASE/Medline. Deze records zijn dus ook goed vindbaar voor ervaren onderzoekers die gewend zijn dit vocabulaire te gebruiken. Binnen Scopus is dit vocabulaire niet apart per type term doorzoekbaar. De termen maken deel uit van één keywords-veld. De Scopus-records die niet uit een eigen onderliggende database afkomstig zijn hebben vaak alleen auteurstrefwoorden.

Tabel 4.9 Verschillen tussen aanwezige velden voor 1 artikel in parallelle records van 7 bibliografische databases.

Etienne, S (2003) <i>Ecological impact in data-poor systems: a case study on metapopulation persistence in geselecteerde databases</i> , 20060414							
velden waarin er verschil is tussen de databases	Biosis silverpl.	Embase silverpl.	Geobase silverpl.	Medline silverpl.	Scopus	WoS	Springer
instelling, adres	1	1	1	1	alle*	alle**	3
DOI / View at publisher	N	N	N	N	J	J	J
documenttype	J	J	N	J	J	J	N
copyright	N	J	J	J	J	J	N
taal	J	J	J	J	J	J	N
aantal literatuurverwijzingen	N	N	J	N	J	J	N
literatuurverwijzingen	N	N	N	N	J	J	J
totaal aantal keywords	44	14	10	6	30	15	5
opmerkingen keywords	hoog aantal door biologische soorten-thesaurus				5 author en 25 uit 4 andere bronnen (Compendex / Embase / Geobase / Medline)	5 author en 15 gegenereerd uit referenties (Key-words-plus)	

bron: naar een idee van Van Laarhoven (UB Groningen)

* voor oudere jaren vaak maar 1-3 adressen; ** auteurs niet 1-op-1 aan affiliaties gekoppeld en voor oudere jaren vaak maar 1-3 adressen

WoS aan de andere kant heeft naast de auteurstrefwoorden geen termen uit gecontroleerd vocabulaire of vakmatige wordsystemen, maar wel overal de 'keywords plus', termen die automatisch gegenereerd worden op basis van veelvoorkomende woorden en concepten in de titels van literatuur waarnaar bij een artikel verwezen wordt. Uit een vergelijkende studie (Qin 2000) blijkt voorzichtig dat zowel de keywords-plus als de 'controlled vocabulary' (bv. thesaurustermen) doorgaans de belangrijkste concepten in een artikel bevatten en dat aanvullende termen van beide hun eigen waarde hebben.

Een ander belangrijk onderdeel van de records is het abstract. Hier wordt standaard dor WoS en Scopus ook in gezocht. Uiteraard zijn abstracts bij bepaalde publicatievormen niet beschikbaar, maar ook bij sommige tijdschriften zijn er geen records beschikbaar in Scopus: soms omdat ze er niet zijn (bij de vakbladen, 'industry journals'), soms omdat Scopus ze niet heeft kunnen opnemen. Jasco (2006) schat dat 20 van de 28 miljoen records abstracts bevatten. Eigen tests (zoeken op *a/the/an* in het abstract) laat inderdaad zien dat minimaal 70% abstracts heeft: 19,48 miljoen van het totaal van 27,97 miljoen records (juni 2006).

Vergeleken met vakbibliografieën ontbreken in Scopus en WoS vaak specifieke vakgerelateerde ontsluitingsvelden. Voorbeelden van gegevens die helemaal ontbreken, zijn:

- geografische coördinaten, zoals in Georef;
- molecuulstructuren, zoals in Chemical Abstracts (in UU via SciFinder Scholar); koppeling met Crossfire-Beilstein is overigens wel mogelijk;
- leeftijdscategorie, "population group" en gebruikte methodologie, zoals in PsycINFO.

Sommige gegevens zijn wel aanwezig, maar dan opgenomen in het bredere keywords-veld en daardoor niet apart doorzoekbaar en ook niet zo consequent aanwezig als in de genoemde vakbibliografieën. Dit geldt voor:

- geografische locaties (zoals gebruikt in Geography/Geobase, Econlit en CAB);
- biologische soorten (zoals gebruikt in Geography/Geobase en CAB).

Andere gegevens zijn wel aanwezig en apart doorzoekbaar, zoals CAS registry numbers voor chemische verbindingen (ook gebruikt in Chemical Abstracts, Pubmed, EMBase en CAB) en genetische sequenties.

4.6 Citatiegegevens: de dekking van “citerende artikelen”

Naast de 28 miljoen records bevat Scopus volgens eigen opgave ook nog eens 245 miljoen literatuurverwijzingen uit die records. Een deel daarvan verwijst terug naar een van 27 miljoen records. Echter, alleen aan records vanaf 1996 zijn gegevens over inkomende citaties gelinked. Jascó (2006) schat dat het uiteindelijk 9,5 miljoen records betreft die zo ‘citation enhanced’ zijn.

Ten behoeve van gebruik van Scopus als citatie-index is het belangrijk inzicht te hebben in de prestaties van het systeem bij het vinden van ‘citerende artikelen’. Gezien de aard van het citatie-zoeken, is een vergelijking met andere citatie-indexen de enige manier om hierover kwantitatieve gegevens te krijgen. Uit eerder onderzoek (Bakkalbassi 2006) bleek reeds dat verschillende indexen voor verschillende vakgebieden uiteenlopen in citatie-kwantiteit. Een extra complicatie hierbij is dat hierbij een vergelijking van alleen getallen niet volstaat. Immers als Scopus bij een artikel 40 citaties vindt en een andere citatie-index ook, dan hoeven dat nog niet dezelfde 40 artikelen te zijn. Daarom is een wat uitgebreider onderzoek gedaan waarin de citaties die met verschillende systemen werden gevonden ook op het niveau van individuele artikelen zijn vergeleken. Gezien de complexiteit van een dergelijke vergelijking is volstaan met een wat beperkte steekproef.

Het onderzoek is als volgt uitgevoerd:

Vergeleken citatie-indexen: Scopus, Web of Science, Google Scholar.

Aantal vergeleken uitgangsdokument: 64 artikelen die in alle drie systemen zelf als artikel voorkwamen; om de resultaten voldoende vergelijkbaar te maken is dus niet gekeken naar artikelen die in Web of Science alleen maar geciteerd worden, maar daarin niet zelf voorkomen.

Verdere randvoorwaarden hierbij:

- gezien de dekking in de tijd van Scopus als citatie-index, is gekozen voor 32 artikelen uit 1995 en 32 artikelen uit 2000;
- uit elk van de 18 in UBU-vakgebieden zijn 4 artikelen geselecteerd, 2 uit 1995 en 2 uit 2000; omdat voor de vakgebieden Godgeleerdheid en Wijsbegeerte niet voldoende artikelen gevonden konden worden die aan de verdere randvoorwaarden voldeden, zijn die beide vakgebieden weggelaten, zodat 16 vakgebieden overbleven;
- de titels van tijdschriften waaruit de onderzochte artikelen afkomstig waren, zijn naar evenredigheid over het alfabet gespreid;
- in verband met de handmatige vergelijking op titelniveau van de citerende artikelen is gekozen voor 32 artikelen die niet vaker dan 50 keer geciteerd werden in Scopus en 32 artikelen die niet vaker dan 50 keer geciteerd werden in Web of Science;
- om ook een voldoende aantal citaties te kunnen vergelijken, zijn op dezelfde manier artikelen geselecteerd die ten minste 30 keer geciteerd werden.

De op deze wijze verzamelde kwantitatieve gegevens zijn samengevat (tabel 4.10 & 4.11). Gegeven zijn de totale aantallen citerende artikelen per categorie. Behalve een uitsplitsing naar de 16 vakgebieden, zijn ook totalen per publicatiejaar en per brede groep alfa-beta-gamma gegeven.

Tabel 4.10 Citaties van geselecteerde artikelen in Scopus, WoS en Google Scholar, totaal en per vakgebied, met overlapgegevens, april 2006.

aantallen citerende artikelen	cumu- latief totaal	Scopus	uniek Scopus	WoS	uniek WoS	Google Scholar	uniek Google	overlap S-W	overlap S-G	overlap W-G	overlap S-W-G
alles	4135	2733	242	2581	221	2671	1120	2301	1492	1360	1301
1995	2063	1372	124	1310	120	1273	543	1161	702	643	615
2000	2072	1361	118	1271	101	1398	577	1140	790	717	686
"beta"	2489	1933	177	1787	88	1501	428	1658	1033	975	935
"gamma"	1437	667	54	657	118	1038	638	528	388	314	302
"alfa"	209	133	11	137	15	132	54	115	71	71	64
aardwet.	216	169	9	163	5	128	39	154	86	83	80
biologie	211	188	9	177	6	154	16	170	137	129	128
diergen.	237	171	8	172	6	145	56	162	85	88	84
economie	514	205	17	174	8	435	299	164	134	112	110
farmacie	229	182	22	154	6	139	41	148	98	86	86
geneesk.	263	202	29	193	21	156	34	166	116	115	109
landbouw	204	164	14	149	4	125	32	141	89	84	80
letteren	209	133	11	137	15	132	54	115	71	71	64
milieuk.	207	179	16	166	8	116	17	155	96	91	88
nat.&ster.	204	175	7	181	13	64	13	165	48	48	45
rechtsgel.	263	98	5	168	76	145	82	85	56	55	48
scheikunde	185	173	4	174	5	98	7	169	91	91	91
soc.geogr.	317	170	15	135	10	225	136	126	89	59	59
soc.weten	343	194	17	180	24	233	121	153	109	88	85
techniek	187	151	23	133	7	108	24	121	79	77	72
wisk.&inf.	346	179	36	125	7	268	149	107	108	83	72

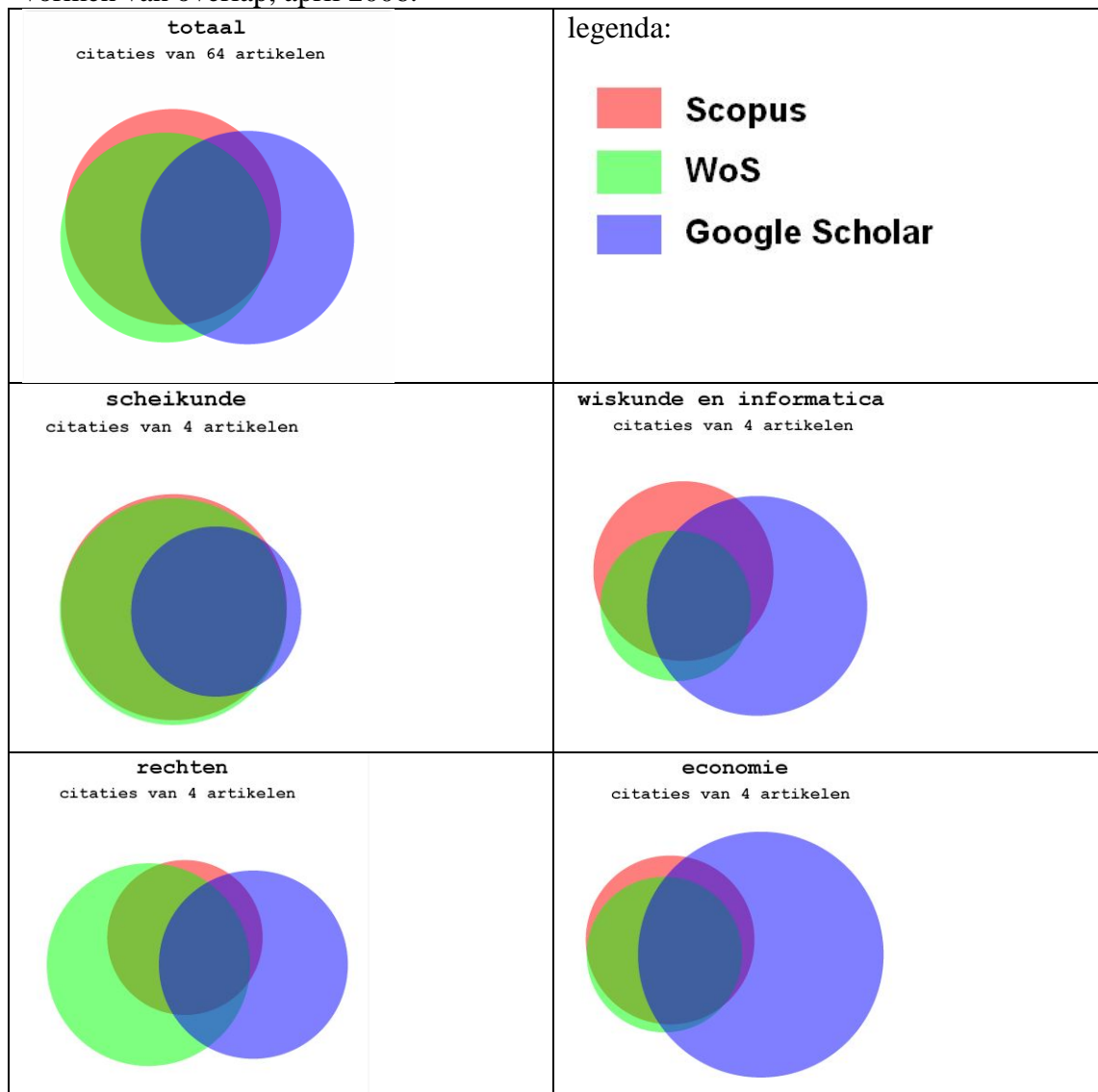
Tabel 4.11 Citaties van geselecteerde artikelen in Scopus, WoS en Google Scholar, totaal en per vakgebied, met overlapgegevens, index (Scopus=1.00), april 2006.

idem, ge- normeerd op Scopus	cumu- latief totaal	Scopus	uniek Scopus	WoS	uniek WoS	Google Scholar	uniek Google	overlap S-W	overlap S-G	overlap W-G	overlap S-W-G
alles	1.51	1.00	0.09	0.94	0.08	0.98	0.41	0.84	0.55	0.50	0.48
1995	1.50	1.00	0.09	0.95	0.09	0.93	0.40	0.85	0.51	0.47	0.45
2000	1.52	1.00	0.09	0.93	0.07	1.03	0.42	0.84	0.58	0.53	0.50
"beta"	1.29	1.00	0.09	0.92	0.05	0.78	0.22	0.86	0.53	0.50	0.48
"gamma"	2.15	1.00	0.08	0.99	0.18	1.56	0.96	0.79	0.58	0.47	0.45
"alfa"	1.57	1.00	0.08	1.03	0.11	0.99	0.41	0.86	0.53	0.53	0.48
aardwet.	1.28	1.00	0.05	0.96	0.03	0.76	0.23	0.91	0.51	0.49	0.47
biologie	1.12	1.00	0.05	0.94	0.03	0.82	0.09	0.90	0.73	0.69	0.68
diergen.	1.39	1.00	0.05	1.01	0.04	0.85	0.33	0.95	0.50	0.51	0.49
economie.	2.51	1.00	0.08	0.85	0.04	2.12	1.46	0.80	0.65	0.55	0.54
farmacie	1.26	1.00	0.12	0.85	0.03	0.76	0.23	0.81	0.54	0.47	0.47
geneesk.	1.30	1.00	0.14	0.96	0.10	0.77	0.17	0.82	0.57	0.57	0.54
landbouw	1.24	1.00	0.09	0.91	0.02	0.76	0.20	0.86	0.54	0.51	0.49
letteren	1.57	1.00	0.08	1.03	0.11	0.99	0.41	0.86	0.53	0.53	0.48
milieuk.	1.16	1.00	0.09	0.93	0.04	0.65	0.09	0.87	0.54	0.51	0.49
nat.&ster	1.17	1.00	0.04	1.03	0.07	0.37	0.07	0.94	0.27	0.27	0.26
rechtsgel.	2.68	1.00	0.05	1.71	0.78	1.48	0.84	0.87	0.57	0.56	0.49
scheikunde	1.07	1.00	0.02	1.01	0.03	0.57	0.04	0.98	0.53	0.53	0.53
soc.geogr.	1.86	1.00	0.09	0.79	0.06	1.32	0.80	0.74	0.52	0.35	0.35
soc.weten	1.77	1.00	0.09	0.93	0.12	1.20	0.62	0.79	0.56	0.45	0.44
techniek	1.24	1.00	0.15	0.88	0.05	0.72	0.16	0.80	0.52	0.51	0.48
wisk.&inf.	1.93	1.00	0.20	0.70	0.04	1.50	0.83	0.60	0.60	0.46	0.40

De algemene conclusie van het deelonderzoek over de citatiegegevens luidt dat de verschillen in dit soort dekking tussen Scopus en Web of Science in het algemeen maar klein zijn; die tussen Google Scholar en deze beide commerciële citatie-indexen waren aanzienlijk groter. Meer in detail kijkend kan - met enige voorzichtigheid vanwege de relatief kleine steekproef - nog het volgende worden geconcludeerd:

- Er is geen verschil waargenomen tussen de oudere (1995) en de meer recente publicaties (2000).
- Bij de meeste vakgebieden ligt de overlap tussen Scopus en WoS in citerende artikelen tussen de 80 en 90% (ten opzichte van de aantallen Scopus-citaties); bij Diergeneeskunde, Natuur- & Sterrenkunde en (vooral) Scheikunde ligt de overlap zelfs nog hoger; bij Sociale Wetenschappen en Sociale Geografie & Planologie ligt de overlap wat lager, bij Wiskunde & Informatica zelfs aanzienlijk lager.
- Alleen voor Rechtsgeleerdheid is het aantal gevonden citaties in WoS aanzienlijk hoger dan dat in Scopus. Gerelateerd aan de WoS-totalen, zou de overlap voor dat vakgebied dus ook veel kleiner geweest zijn.
- Voor Wiskunde & Informatica en in mindere mate voor Sociale Geografie & Planologie is het aantal gevonden citaties in Scopus significant hoger dan in WoS.
- Ten aanzien van Google Scholar - hoewel minder relevant voor de huidige vergelijking - kan geconcludeerd worden dat de aanzienlijk geringere overlap met de resultaten uit de commerciële databases vooral wordt veroorzaakt door de veel grotere aantallen publicaties in andere talen dan Engels en andere documenttypes die in Google Scholar zijn opgenomen; in het bijzonder voor Economie, Rechtsgeleerdheid, Sociale Geografie & Planologie en Wiskunde & Informatica liggen de gevonden aantallen citaties aanzienlijk hoger dan bij Scopus; voor de meeste (andere) bètavakgebieden liggen die aantallen juist significant lager.
- De resultaten van deze citatiesteekproeven voor de individuele vakgebieden komen redelijk overeen met de dekkingsgegevens uit hoofdstuk 4.1.3.

Figuur 4.10 Overlap van citaties naar alle geselecteerde artikelen tussen Scopus, WoS en Google Scholar en voor artikelen uit enkele geselecteerde vakgebieden met uiteenlopende vormen van overlap, april 2006.



bron: eigen onderzoek

5 Zoekfunctionaliteit, interface, snelheid en gebruiksgemak

5.1 Zoekfunctionaliteit

Scopus en WoS hebben veel overeenkomsten, ook in functionaliteit. Vaak verschilt alleen de vormgeving of plaatsing van bepaalde items op het scherm of binnen de site. Hoewel dat voor individuen kan schelen voor de look and feel, is het lastig om er een algemeen oordeel over te geven. Daarom richten we ons hier op die functionaliteit die wel verschilt, dat wil zeggen: zaken die in het ene systeem wel mogelijk zijn en in het andere niet (tabel 5.1).

Tabel 5.1. 'Harde' functionaliteitsverschillen (stand juni 2006)

Wel mogelijk in Scopus, niet in WoS	Wel mogelijk in WoS, niet in Scopus
- default refine (parametrisch zoekresultaat)	- uitsplitsing naar land, stad, affiliatie
- citatietabel (Citation Tracker)	- koppeling naar Journal Citation Reports
- zoeken op 'alles'	- doorverwijzing naar Crosssearch bij 0 results
- zoeken op casregnummers (chemische stoffen)	- doorverwijzing naar Current Web Contents
- integratie met Beilstein (molecuulstructuren)	- trunceren binnen phrase: neural network*
- nabijheidszoeken met PRE en W	
- zoeken in alleen ABS of KEYW of AUTHKEY	
- zoeken op genetische sequenties	
- gebruik Author Identifier (wel verwacht in WoS)	
- koppeling met Scirus	
- automatische stemming (meervoudsvormen)	
- UBUlink direct in lijstweergave references	
- tijdschriftenlijst browseable per vakgebied	
- tijdschriftenlijst doorzoekbaar op uitgever	

Scopus is iets veelzijdiger dan Web of Science. Met name de standaard refine-balk is een pluspunt, niet alleen bij het zoeken, maar ook (voor studenten) om grip te krijgen op een onderzoeksveld via uitsplitsing van het zoekresultaat naar brontijdschriften, publicatiejaren, auteurs en overlap van vakgebieden die je bij een zoekactie krijgt. Een tweede groot pluspunt van Scopus is het gemak van het creëren van overzichtelijke citatietabellen van een auteur of onderwerp. Wel is deze mogelijkheid te zeer verstopt.

De Author Identifier van Scopus tenslotte is een al lang gekoesterde wens voor het groeperen van gelijke auteurs en het scheiden van verschillende. Met een algoritme waarin o.a. affiliatie, co-auteurs en citaties een rol spelen worden verschillende notaties van de naam van één auteur gegroepeerd en verschillende auteurs met exact dezelfde naam en voorletters gescheiden. Elke auteur krijgt een uniek ID. Het systeem werkt, maar kan niet (nog) alle auteursnamen groeperen. Wel is hierbij de mogelijkheid om om simpele wijze feedback te leveren. Een bijkomend voordeel van de author identifier is dat bij oudere artikelen waar niet altijd voor elke auteur een affiliatie bekend is, nu via een link naar de auteursgegevens toch affiliaties opgezocht kunnen worden. Overigens heeft Web of Science aangekondigd iets soortgelijks als de Author Identifier te gaan implementeren.

Web of Science biedt iets meer mogelijkheden voor gevorderde citatieanalyse van complete organisaties, vooral de mogelijkheden het zoekresultaten uit te splitsen naar affiliatie: de organisatie waar de auteur werkzaam is en de stad en het land waar deze organisatie is

gevestigd. Beide bedrijven zijn in deze periode doende hun zoeksystemen te verbeteren. Dat lijkt in elk geval al een voordeel van de concurrentie die met Scopus gekomen is.

5.2 Interface, snelheid en gebruiksgemak

De snelheid van zoeksystemen is een belangrijk element van de 'feel'. Het bepaalt deels of mensen een systeem prettig vinden werken. De snelheid is uiteraard van veel zaken afhankelijk: de server, de verbinding, de browser, andere taken die de PC aan het uitvoeren is en de specifieke taak die aan het zoekstelsel wordt gegeven. In een simpele test hebben we voor vier zoeksystemen zoveel mogelijk van deze zaken gelijk gehouden (tabel 5.2). Zowel Scopus als WoS hebben complexe interfaces die vrij langzaam opbouwen, zowel bij starten van het systeem als bij zoekacties. Het Utrechtse systeem Omega is zeer snel, maar Google Scholar blijkt bij de meeste acties het snelste te werken. Dit zal zeker een rol spelen in het enthousiasme van velen voor deze zoekmachine. Wel moet natuurlijk worden bedacht dat de aard van de informatie die wordt getoond en de geboden functionaliteit van Scopus en WoS in de fase van het selecteren uit de zoekresultaten tijdswinst kan opleveren ten opzichte van Omega en Google Scholar. De onderlinge verschillen tussen Scopus en WoS zijn erg klein.

Tabel 5.2. Snelheid van brede artikelzoeksystemen, juni 2006, paginaopbouw in seconden.

	Google Scholar	Omega zoekmachine	Scopus	Web of Science *
Opbouwen zoekscherm na ingeven URL	<1	5	7	10 **
"lymphocyte development" AND thymus	1	2 ***	8	13
trauma AND psychopathology	1	1	10	8
embryogenesis AND bovine	2	1	8	9

NB Uitgevoerd met MSIE 6.0, alle overige programma's uitgeschakeld, binnen 5 minuten na elkaar. Cijfers zijn gemiddelde van twee gelijke zoekacties op verschillende dagen. Zoekacties in standaard velden, aantal resultaten per pagina op 100.

* Aantal resultaten per pagina 50 (maximum); ** Inclusief 1 maal doorklikken van Portal naar zoekscherm; *** Omega ondersteunt phrase search niet.

Naast de snelheid wordt het gebruiksgemak ook bepaald door de helderheid van de navigatie en layout. Deze hebben we binnen het kader van dit onderzoek niet zelf getest. Uit beoordeling door gebruikers (hoofdstuk 6) komt wel naar voren dat men doorgaans de inrichting van Scopus iets duidelijker vinden dan die van WoS. Gezien de snelle ontwikkeling van de user interface van beide systemen lijkt dit niet een breekpunt in een eventuele keuze te moeten zijn.

5.3 Vakgebiedsindeling

Bij multidisciplinaire bestanden is het bij sommige acties prettig te kunnen inzoomen op documenten uit een bepaald vakgebied. Als men heel zeker is van de zaak kan dit direct bij

een zoekactie als beperkende voorwaarde worden meegegeven, maar doorgaans zal men dit pas in tweede instantie doen om een zoekresultaat te beperken tot een bepaald vakmatig perspectief. Zo kan men artikelen over aspecten van genetica willen beperken tot die geschreven in het vakgebied van de sociale wetenschappen in plaats van de medische of biologische. Of men wil aspecten van transport bestuderen vanuit economisch in plaats van technisch perspectief. In dat soort gevallen is het belangrijk te weten wat de grenzen van de vakgebiedsindeling van een bestand zijn en hoe die zich verhouden tot reeds bekende indelingen. Scopus heeft een indeling in 27 vakgebieden, die weer geaggregeerd zijn tot 4 richtingen (health, life, physical en social). Na een zoekactie is in de refine tool direct te zien in welke vakgebieden het onderwerp blijkbaar bestudeerd wordt.

Figuur 5.1 Vakgebiedsindelingen van Omega en Scopus gekoppeld.

Scopus*	UBU
P: earth & planetary sciences	Aardwetenschappen
H/L/P/S: multidisciplinary	Algemeen
L: agricultural and biological sciences	Biologie
L: biochemistry, genetics and molecular biology	
L: immunology and microbiology	
H: veterinary	Diergeneeskunde
S: economics, econometrics and finance	Economie
S: business, management and accounting	
L: pharmacology, toxicology and pharmaceuticals	Farmacie
H: medicine	Geneeskunde
H: neuroscience	
H: nursing	
H: health professions	
H: dentistry	Godgeleerdheid
S: arts and humanities	Landbouwwetenschappen
P: environmental science	Letteren
P: energy	Milieukunde
P: physics and astronomy	Natuur- & Sterrenkunde
P: materials science	Rechtsgeleerdheid
P: chemistry	Scheikunde
P: chemical engineering	Sociale Geografie & Planologie
S: social sciences	Sociale Wetenschappen
S: psychology	Techniek
P: engineering	Wijsbegeerte
P: mathematics	Wiskunde en Informatica
P: computer science	
P: decision sciences	

* H=health sciences; L=life sciences; P=physical sciences; S=social sciences

De 27 vakgebieden zijn gevormd door tijdschriften (als geheel) in te delen, waarbij dus geldt dat alle artikelen uit dat tijdschrift in het vakgebied vallen. Wel is het zo dat een tijdschrift bij meer vakgebieden ingedeeld kan zijn. Dit geldt voor algemene wetenschappelijke tijdschriften (bv. Nature), maar ook voor meer gespecialiseerde. In totaal is er zo in Scopus 30% verdubbeling van opgenomen records. Dit is duidelijk te zien als men na een zoekactie inzoomt op een vakgebied met behulp van de refine tool. Dan blijkt dat vakgebied toch weer records van andere vakgebieden te bevatten.

Omdat er geen algemeen geldende indeling van wetenschappelijke vakgebieden is, en gebruikte indelingen nooit één op één op elkaar passen, is het goed om te zien hoe de in de UBU gehanteerde (Omega-)indeling zich verhoudt tot die van Scopus (figuur 5.1).

De meeste UU-vakgebieden zijn grover dan die van Scopus, waardoor selecteren van relevante vakgebieden in Scopus vrij eenvoudig is en goede filtermogelijkheden oplevert waarmee alleen relevant materiaal binnen de selectie blijft zonder te veel relevants buiten te sluiten. Alleen voor de sociale geografie & planologie geldt dat de Scopus indeling waarin dit materiaal valt veel grover is: social sciences. Sociaal-geografen en planologen hebben dus relatief weinig aan de Scopus-indeling.

6 Beoordeling door gebruikers

6.1 Interviews met veelgebruikende onderzoekers

Om inzicht te krijgen in hoe veeleisende onderzoekers citatiebestanden gebruiken hebben we gesprekken gevoerd met 22 personen (zie tabel 6.1). De puntsgewijze verslagen van de gesprekken staan in bijlage III. Bij de gesprekken was behalve een lid van de Scopus onderzoeksgroep ook de betreffende UBU-vakspecialist aanwezig.

Tabel 6.1 Geïnterviewde veelgebruikers

Faculteit	Aantal geïnterviewde onderzoekers
Bètafaculteit	8
Diergeneeskunde	3
Geesteswetenschappen	0
Geneeskunde	2
Geowetenschappen	4
Letteren	0
Rechten	0
Sociale Wetenschappen	5

Het algemene beeld in de interviews met veelgebruikers is er een van grote verscheidenheid, naar gebruikte bestanden zowel als naar gebruikswijze van bestanden. Sommige onderzoekers zoeken zeer gericht, vooral naar *known items* en op citaties. Anderen zoeken veel meer op inhoudelijke termen en ook inhoudelijk breder, inventariserend. Ook is duidelijk geworden dat her en der in de UU citatiebestanden voor niet vakinhoudelijke doelen worden gebruikt: als verificatie bij Metis-invoer, als hulpmiddel bij planning van onderzoeksprogramma's en als leverancier van primaire gegevens voor onderzoek naar innovativiteit van landen en regio's of op het terrein van de wetenschapssociologie.

Vrijwel unaniem wordt Scopus geprezen om de interface, heldere schermopbouw en navigatie en de standaard zichtbare refine-opties (het parametrisch zoeken, dat overigens ook in Omega wordt geïmplementeerd). Daarbij moet worden aangetekend dat de meerderheid dat ook in vergelijkende zin een pluspunt vindt ten opzichte van WoS, waar men vaker "de weg kwijt is" of "omslachtige handelingen moet verrichten", vooral bij de *cited ref* zoekacties.

Opvallend is dat vrijwel niemand ondanks de geprezen interface eigener beweging de mogelijkheid van de citation overview had gevonden. Die tonende is iedereen zeer enthousiast, vooral over de heldere tabelgewijze presentatie van citaties van artikelen van een specifieke auteur. Die presentatie is na de interviews in de Spring release 2006 van Scopus overigens verder verbeterd met sorteer- en doorklikopties. Een minderheid zegt de mogelijkheid citatieoverzichten op onderwerp (na een zoekactie|select all|save to list|my list|select all|citation overview) te zullen gebruiken, zelfs als het makkelijker en sneller zou kunnen.

Het gebruik van WoS is intensief, vooral vanwege het citatiezoeken, waarop ISI tot voor kort vrijwel een monopolie had. Zeer verdeeld is de mate waarin WoS als belangrijkste inhoudelijk bestand wordt gebruikt dan wel als een van de ingangen naast andere brede

bestanden (Omega, Picarta, Google Scholar), vakindexen (bijvoorbeeld Pubmed, PsycInfo, Georef, Econlit, CAB-Abstracts) en uitgeversplatforms (bv. Science Direct). Opvallend is dat een aantal onderzoekers aangeeft de betreffende vakindex in hun vakgebied in het geheel niet te gebruiken. Doorgaans heeft men dan de indruk met WoS/Scopus, Omega en Google Scholar voldoende recall te krijgen.

Bij gerichte vragen over hoe WoS en Scopus zich tot elkaar verhouden is men geneigd voor Scopus te kiezen qua gebruiksgemak, maar liggen de kaarten verdeeld als het gaat om dekking. Daarover is bij een niet onaanzienlijk aantal personen onzekerheid, onbekendheid en ongerustheid. Deels kan de onzekerheid tijdens het gesprek worden weggenomen; dan bleek dat bij een eerlijke vergelijking de dekking van Scopus beter was of min of meer gelijk. Door tijdschriftenlijsten te laten zien konden de geïnterviewden een indruk krijgen van het dekkingsverschil. Dat leidde zelden tot een beeld van grote omissies in Scopus t.o.v. WoS. Er blijft wel ongerustheid over de dekking bij de maatschappijwetenschappen, met name voor de periode voor 1996.

De web results en patenten worden door een minderheid wel interessant gevonden, maar dan toch van ondergeschikt belang geacht, omdat deze ook buiten Scopus beschikbaar zijn. Geen van de geïnterviewden zegt geïntegreerde weergave van zoekresultaten van het web (Scirus) en Scopus zelf prettig te vinden (hoewel daar niet systematisch naar gevraagd is).

6.2 Webenquête onder gebruikers

6.2.1 Conclusies

Uit een analyse van een gebruikersonderzoek onder 81 personen zijn een aantal interessante conclusies te trekken. De belangrijkste zijn:

- Google Scholar heeft een grotere naamsbekendheid dan Scopus of Web of Science.
- Scopus wordt vooral gebruikt door studenten en wetenschappers uit de bèta en gamma hoek, alfa gebruikt meer Google Scholar; studenten gebruiken relatief gezien meer Google Scholar, AIO/wetenschappers juist meer Web of Science.
- Respondenten vinden dat Scopus de beste treffers oplevert, gevolgd door Web of Science; over Google Scholar is slechts 15 % echt tevreden. Respondenten met een bèta achtergrond zijn allemaal (!) tevreden over de kwaliteit van de treffers in Scopus.
- Bij Google Scholar zijn de meningen over de kwaliteit van de treffers erg verdeeld: van de AIO/wetenschappers is 65% ontevreden, terwijl van de studenten slechts 22% ontevreden is; 22% van de alfa groep is ontevreden, van de bèta groep is dat 61%.
- 84% van de respondenten vindt Scopus prettig of redelijk makkelijk in gebruik; Google Scholar scoort 80%, Web of Science 61%. Slechts 50% van de studenten vinden Web of Science prettig of redelijk makkelijk in gebruik, van de AIO/wetenschappers is dat 76%.

Scopus wordt dus beduidend beter beoordeeld dan Web of Science, zowel wat betreft de gebruikersvriendelijkheid als de kwaliteit van de treffers.

6.2.2 Doel van het onderzoek

De UBU onderzoekt of het bestand Scopus van Elsevier aangeschaft moet worden. Dit bestand wordt getest door middel van een trial. Naast andere onderzoeken naar de dekking en functionaliteiten van Scopus, is dit onderzoek gehouden om meer te weten te komen van de voorkeuren van de gebruikers, vooral in vergelijking tot Web of Science.

Het doel van dit onderzoek is dus om een statistisch onderbouwd beeld te krijgen van de volgende zaken:

- Hoe vaak gebruikt men Scopus, vergeleken met Web of Science en Google Scholar?
- Wat vindt men van de kwaliteit van de zoekresultaten van Scopus en Web of Science?
- Wat vindt men van het gebruikersgemak van Scopus en Web of Science?
- Hoe heeft men gehoord van Scopus?

6.2.3 Opzet

De enquête is uitgevoerd in april en mei 2006 met een online tool van www.surveymonkey.com. Hiermee is eenvoudig een digitale enquête te maken, waarvan de resultaten direct kunnen worden opgevraagd. Er is een nieuwsbericht geplaatst op diverse bibliotheekwebsites. Ook is er aandacht op het onderzoek gevestigd in een aantal digitale nieuwsbrieven.

Er hebben uiteindelijk 94 personen gereageerd. Dertien reacties zijn verwijderd omdat ze niet verder kwamen dan vraag 2; uiteindelijk zijn er dus 81 ingevulde vragenlijsten geanalyseerd.

Om het aantal respondenten te vergroten is er voor gezorgd dat de vragenlijst in ongeveer tien minuten is in te vullen. Achteraf was te zien dat bijna iedereen binnen de tien minuten klaar was.

6.2.4 Uitkomst digitale vragenlijst

1. Geef aan wat op u van toepassing is:

Bachelor Student	19	23%
Master student	17	21%
AIO/Promovendus	17	21%
Wetenschapper	19	23%
Overig Universitair medewerker	6	7%
Niet verbonden aan de universiteit	3	4%
Totaal	81	100%

Alle relevante groepen zijn voldoende vertegenwoordigd in dit onderzoek. In de verdere analyse zal deze groep worden samengevat in studenten, aio/wetenschapper en overig.

2. Wat is uw vakgebied?

Biologie	3	4%
Farmacie	8	10%
Geneeskunde	2	2%
Geowetenschappen	7	9%
Godgeleerdheid	1	1%
Letteren	10	12%
Rechtsgeleerdheid / Economie	4	5%
Scheikunde	12	15%
Sociale wetenschappen	27	33%
Wiskunde en Informatica	1	1%

Overig	6	7%
Totaal	81	100%

Sociale wetenschap en scheikunde zijn oververtegenwoordigd; deze groepen zijn dan ook actiever benaderd om mee te doen aan dit onderzoek. In de verdere analyse zal deze groep worden samengevat in alfa, bèta (incl. medisch), en gamma. Bij overige werden genoemd: biomedische wetenschappen (neurowetenschappen), milieuwetenschappen, onderwijskunde, toxicologie en UBU (2x).

3. Hoe vaak gebruikt u onderstaande zoekmachines voor wetenschappelijke literatuur?

	Scopus		Web of Science		Google Scholar	
Meerdere keren per week	20	25%	12	15%	15	19%
Een paar keer per maand	19	23%	20	25%	22	27%
Een paar keer per jaar	8	10%	11	14%	10	12%
Zelden of nooit	11	14%	18	22%	22	27%
Ik ken deze zoekmachine niet	23	28%	20	25%	12	15%
Totaal	81	100%	81	100%	81	100%

Omdat het onderzoek over Scopus ging, is het aannemelijk dat Scopus gebruikers oververtegenwoordigd zijn in dit onderzoek. Er valt dus niet te concluderen dat Scopus meer wordt gebruikt; dat wordt ook door de statistieken niet ondersteund. Opvallend is wel dat Google Scholar door relatief weinig mensen niet wordt gekend (de beste score van de drie zoekmachines) en vaker wordt gebruikt dan Web of Science.

In verdere analyse zal deze groep (per zoekmachine) worden samengevat in gebruikers (die de zoekmachine minstens een paar keer per jaar gebruiken) en niet-gebruikers (de rest).

Scopus wordt relatief meer gebruikt door bèta en gamma en minder door alfa. De Web of Science gebruikers zijn relatief veel AIO/wetenschappers en beduidend minder studenten. Google scholar wordt juist door meer studenten gebruikt en ook relatief meer door alfa.

4. We willen u vragen om een voor u relevante zoekvraag te stellen in Scopus (link via MyUU), Web of Science (link via MyUU) en Google Scholar, dus dezelfde vraag in alle drie de zoekmachines. Kunt u daarna aangeven wat u vindt van de kwaliteit van de treffers die u krijgt in deze zoekmachine?

	Scopus		Web of Science		Google Scholar	
Hoge kwaliteit/relevantie	25	49%	13	27%	8	15%
Redelijke kwaliteit/relevantie	19	37%	17	35%	21	40%
Matige kwaliteit/relevantie	5	10%	10	20%	20	38%
Slechte kwaliteit/relevantie	2	4%	9	18%	3	6%
Ken zoekmachine niet of gebruik deze niet	9		11		8	
Totaal	60	100%	60	100%	60	100%

De percentages zijn berekend zonder de groep die de zoekmachine niet kent of gebruikt mee te tellen. Opvallend is dat ook 21 respondenten bij deze vraag zijn afgehaakt, blijkbaar vonden ze deze vraag te veel werk.

Duidelijk is dat Scopus substantieel beter scoort, gevolgd door Web of Science en daarna Google Scholar.

In verdere analyse is deze groep samengevat in tevreden (die kozen voor hoge of redelijke relevantie) en ontevreden gebruikers (die kozen voor matige of slechte relevantie). De onderverdeling in studenten en AIO/wetenschappers is alleen interessant bij Google Scholar: daar is 65% van de laatste groep ontevreden, terwijl van de studenten slechts 22% ontevreden is. De overige zoekmachines laten geen substantieel verschil zien tussen deze groepen.

De verdeling in alfa, bèta en gamma levert wel veel verschillen op. Bij Scopus is de hele bèta groep tevreden (!), in tegenstelling tot de alfa groep: slechts 63% tevreden. Bij Web of Science is de gamma groep meer tevreden (65%) dan de alfa groep (50%). Google scholar levert het meeste onderscheid op: 78% van de alfa groep is tevreden, van de bèta groep is slechts 39% tevreden (gamma wijkt niet af van het gemiddelde).

Verder is te zien dat gebruikers meer tevreden zijn dan niet-gebruikers, maar dat is natuurlijk niet verbazingwekkend.

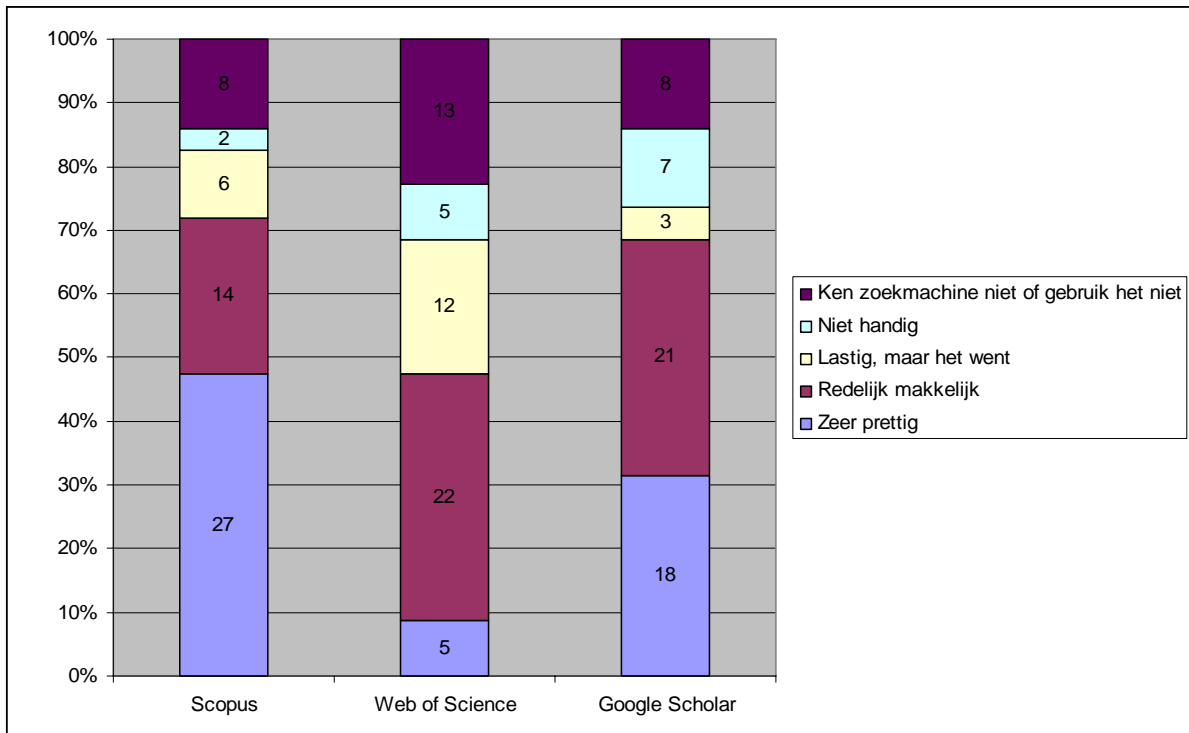
5. Op welke zoektermen heeft u gezocht?

De zoektermen die zijn ingevoerd waren (aanhalingstekens later toegevoegd om termen per respondent te groeperen):

"aggression", "aggressive behavior and information processing", "'American Sign Language' AND agrammatism", "Antichymotrypsin AND Alzheimer Disease", "argument kuhn, d", "auteurs, working memory, language, development", "auteursnaam", "Bayes, diagnosis Psychological", "bijlmermeer", "birth order achievement differences", "bolivia + socialism", "Boschma, R.A.; Competitiveness of Regions from an evolutionary perspective", "cawthon telomere PCR", "chronic low backpain", "computer assisted assessment", "COPD AND collagen", "crowding eye movements", "Down syndrome", "european AND parliament AND portugal", "evidence based practice", "feline hypothyroidism treatment", "fractures bèta blockers", "functie cAMP", "gender-mainstreaming", "greenblatt AND shakespeare", "history university utrecht", "iapp and membrane", "McEnrue, M / Human Resource Development Quarterly / Formal Mentoring Programs", "mergers AND R&D AND relatedness", "microbiologische termen", "netherlands multiculturalism", "orphan drugs and rare disease", "paclitaxel newborn", "phthalocyanine AND amphiphilic", "protein degradation", "rheumatoid arthritis AND moderat* (in titel)", "schizophrenia emotion", "size dependent adsorption", "social learning cumulative cultural evolution", "squaraine", "strafrecht", "'task analysis' 'instructional design'", "te veel op om apart te benoemen", "'University history' en 'F C Donders'".

6. Kunt u aangeven hoe prettig u deze zoekmachines vindt in het gebruik?

	Scopus		Web of Science		Google Scholar	
Zeer prettig	27	55%	5	11%	18	37%
Redelijk makkelijk	14	29%	22	50%	21	43%
Lastig, maar het went	6	12%	12	27%	3	6%
Niet handig	2	4%	5	11%	7	14%
Ken zoekmachine niet of gebruik het niet	8		13		8	16%
Totaal	57	100%	57	100%	57	116%



De percentages zijn berekend zonder de groep die de zoekmachine niet kent of gebruikt mee te tellen. Scopus scoort beter dan Google Scholar en aanzienlijk beter dan Web of Science. Daar stelt 38% dat het gebruik lastig of niet handig is.

Web of Science vindt 50% van de studenten lastig of niet handig, bij AIO/wetenschappers ligt dat juist veel lager: 24%. Bij Google Scholar valt op dat 29% van de bèta groep de zoekmachine lastig of niet handig vindt, en bij gamma is dat slechts 10%.

7. Hoe bent u er op geattendeerd Scopus te gaan gebruiken?

Ik gebruik Scopus niet	12	21%
Via de homepage van mijn vakbibliotheek	5	9%
Via www.library.uu.nl	11	19%
Via mijn docent	6	11%
Via een medestudent of collega	12	21%
anders, namelijk	11	19%
Totaal	57	100%

De antwoorden zijn zeer gevarieerd; zowel websites, docent als collega's zijn blijkbaar belangrijk om mensen op Scopus te wijzen.

Bij 'anders' werd ingevuld:

- via onderwijspersoon van mijn studievereniging
- inleidend vak WAR (sociale geografie)
- nieuwsbrief bibliotheek
- via e-mail van bibliotheek
- via bibliotheekmedewerker
- via vakspecialist (4x)

- workshop literatuur zoeken van de Juridische Bibliotheek
- kon Scopus niet vinden, verstoep bij bibliografieën en naslagwerken

8. Heeft u nog andere opmerkingen over Scopus?

De respondenten is ook gevraagd in vrij tekst opmerkingen te maken over Scopus, Web of Science en Google Scholar. In bijlage IV staan alle opmerkingen genoemd.

7 Samenvatting vakspecifieke resultaten

7.1 Overzicht

Tabel 7.1 Score van Scopus op diverse aspecten van dekking, per UBU-vakgebied

	(a) % dekking UBU- tijdschriften	(b) dekking 1986- 1995 als % van 1996-2005	(c) % dekking UBU- tijdschriften, vgl. met WoS	(d) dekking op basis van zoekacties 1988 en 2005, vgl. met WoS	(e) citaties van selectie artikelen 1995 en 2000, vgl. met WoS
AW	++	++	+	+	0
BIO	++	+	+	+	+
DIER	++	0	+	+	0
ECON	+	-	++	+	++
FARM	++	++	+	+	++
GENEES	++	++	++	0	0
GODG	--	na	--	na	na
LANDB	++	0	+	na	++
LET	--	-	--	na	0
MIL	+	++	++	+	+
NAT&ST	+	-	+	-	0
RECHT	-	na	+		--
SCHEI	+	-	0	+	0
SG&PL	+	+	++	+	++
SW	0	-	+	+	+
TECH	+	+	++	na	++
WIJS	--	na	--	na	na
WI&INF	0	--	+	0	++

(a) --=<20%; -=20-40%; 0=40-60%; +=60-80%; ++=80-100% (tabel 4.4)

(b) --=0-16%; -=16-32%; 0=32-48%; +=48-64%; ++=64-80% (tabel 4.8) BIO en SW gemiddelde van verschillende Scopus vakgebieden, SG&PL inschatting

(c) --=>20%punt minder; -=5-20%punt minder; 0=5%punt minder tot 5%punt meer; +=5-20%punt meer; ++=>20%punt meer (tabel 4.4)

(d) -=Scopus minder zoekacties met >10% meer resultaat dan WoS; 0=Scopus heeft evenveel zoekacties met >10% meer resultaat dan WoS; Scopus heeft bij meer zoekacties >10% meer resultaat dan WoS (tabel 4.6); hier is een driepuntsschaal gebruikt in verband met de grotere rol van toeval in de meetresultaten

(e) --=>10% minder; -=5-10% minder; 0=5% minder tot 5% meer; +=5-10% meer; ++=>10% meer (tabel 4.10)

na=niet beschikbaar

7.2 Opmerkingen per UBU-vakgebied

Binnen dit onderzoek is veel informatie gegenereerd over de waarde Scopus voor specifieke vakgebieden. Die vatten we hieronder samen. Daarbij letten we op de dekkingsindicatoren zoals bijeengezet in tabel 7.1, op beschikbaarheid specifieke keywords, of vakspecifieke functionaliteit, op de bruikbaarheid van de vakgebiedsindeling van Scopus, op opmerkingen van veelgebruikers en op eventuele sterke verschillen binnen het UBU-vakgebied.

De dekking van de humaniora in Scopus is minimaal. Elsevier heeft zich bij de ontwikkeling van de database daar bewust niet op gericht omdat bleek dat de literatuurbehoefte in deze wetenschappen anders ligt (minder tijdschriften, meer boeken; meer behoefte aan ouder

materiaal). Elsevier had ook geen eigen indexen op dit terrein die een basis zouden kunnen vormen. Dit is dus op zich een zeer belangrijk minpunt ten opzichte van de dekking van Web of Science, dat een complete sectie Arts & Humanities heeft. Het gewicht dat we hieraan geven moet afhangen van het gebruik van WoS voor deze vakgebieden en van de wens om deze vakgebieden geïntegreerd met bèta en gamma te kunnen doorzoeken in een citatiebestand.

7.2.1 Aardwetenschappen

Dekking van aardwetenschappen is goed tot zeer goed, ook in vergelijking met WoS. Van de digitale UBU-titels zit 81% in Scopus. De goede dekking is vooral gebaseerd op de onderliggende database Geobase, die vrijwel geheel in Scopus is opgenomen. Het bestand Georef heeft echter, in vergelijking met Scopus, meer niet-tijdschriftmateriaal opgenomen. Het is aan de andere kant een voordeel om via Scopus tegelijk ook tijdschriften op het terrein van de biologie (geobiologie) en techniek (hydrologie) te kunnen doorzoeken. Scopus mist uiteraard de echte thesaurustermen en functionaliteit van Georef, maar heeft wel de descriptoren uit Geobase overgenomen, inclusief de geografische aanduidingen. De Scopus vakgebiedsindeling is goed werkbaar voor aardwetenschappen.

7.2.2 Biologie

Biologie wordt goed gedekt door Scopus, ook in vergelijking met WoS. Meer dan 83% van de digitale UBU-titels zit in Scopus. De onderliggende database Biobase draagt daar aan bij. Voor biologie heeft de UBU overigens geen specifieke vakindexen. Scopus heeft uit Biobase ook de keywords overgenomen. Er is echter geen soortenthesaurus met explode functie. De biologie heeft ook veel aan de miljoenen records uit aanpalende disciplines als geneeskunde en scheikunde. Voor biologie zijn er minimaal drie relevante vakgebiedaanduidingen in Scopus, waarvan er één gedeeld wordt met landbouwwetenschap.

7.2.3 Diergeneeskunde

Diergeneeskunde zit goed in Scopus met 82% van de digitale UBU-titels. Wel is de backfill nog vatbaar voor verbetering. Scopus scoort beter dan WoS wat betreft overlap met CAB-Abstracts. Scopus heeft geen specifieke onderliggende diergeneeskundige database kunnen gebruiken, maar diergeneeskunde profiteert van de goede dekking van geneeskunde en MeSH-termen uit Pubmed die via EMBase in Scopus terechtkomen. Scopus heeft wel een aparte vakgebiedaanduiding 'veterinary'. Ook landbouwwetenschap zit overigens zeer goed in Scopus; voor dat vakgebied wordt een segment van de Scopus-indeling gedeeld met biologie.

7.2.4 Economie

De economie wordt in Scopus niet vanuit een van de onderliggende database gedekt, m.u.v. regionale economie in Geobase. Toch komt economie er met 61% dekking van digitale UBU-titels beter uit dan WoS, en dat geldt ook voor de zoekacties en citatietelling die voor dit rapport zijn gedaan. Wel is er een noodzaak van opname van meer oudere jaargangen. De backfill is met amper 10% zeer matig. Scopus heeft op dit terrein bijna 20% meer citerende artikelen dan WoS volgens ons onderzoek. Voor de toegepaste economie en business studies kan Scopus moeilijk concurreren met de gratis ingangen naar research papers van RepEC-Econpapers en SMEALsearch.

7.2.5 Farmacie

Scopus dekt een groot deel van de farmacietijdschrift die de UBU digitaal in bezit heeft (81%). Ook alle andere indicatoren voor dekking van farmacie zijn zeer positief, ook in vergelijking met WoS. Dit blijkt overigens ook uit tests gedaan door Schneider (2006). Via EMBase profiteert farmacie natuurlijk ook van aanwezigheid van Pubmed/Medline records in Scopus en van de bijbehorende MeSH-termen. Er is een goed op de farmaciebehoefte passende vakgebiedsindeling in Scopus.

7.2.6 Geneeskunde

De dekking van geneeskunde in Scopus is met 83% van ons digitale tijdschriftenbestand op dit gebied uitstekend, zeker in vergelijking met WoS. Alleen Pubmed is nog breder qua dekking. Scopus heeft MeSH-termen en een (bijna) volledige opname van alle records uit EMBase. Ook specifieke functionaliteit als het zoeken op genetische sequenties is aanwezig. Wel is er nog de vraag hoe het relatief sterke verschil in aantal records tussen Scopus en Pubmed voor de recente jaren te verklaren valt. De relatief beperkte score op backfill (tabel 4.8) kan ook worden verklaard door de buitengewoon sterke groei van het aantal medische publicaties in de laatste 10 jaar. Voor Scopus zijn er tenslotte vrij veel gerichte onderwerpsaanduidingen, waaronder ook een aparte voor verplegingswetenschap.

7.2.7 Godgeleerdheid

Als een van de humaniora is de dekking van godgeleerdheid in Scopus zeer minimaal. Er worden geen specifiek theologische tijdschriften gedekt, wel enige op aanpalende terreinen. Zo is er enige dekking op het terrein van bio-ethiek (bioethic*=6600, waarschijnlijk vooral uit Medline) maar verder nauwelijks (theolog*=3600). Voor godgeleerdheid heeft het bestand zeer beperkte waarde.

7.2.8 Letteren

Het brede terrein van de letteren (talen, kunst, geschiedenis) heeft een zeer beperkte dekking in Scopus, een bewuste keuze van de producent. Er is wel redelijke dekking op het gebied van akoestiek, taaltechnologie, computerlinguïstiek e.d. (waarschijnlijk vooral uit Compendex), maar verder is Scopus niet interessant voor letteren.

7.2.9 Milieukunde

De milieukunde is redelijk goed gedekt met 70% van de UBU-titels. Deze dekking gaat terug op de onderliggende databases Geobase, Biobase en Compendex. De backfill is dan ook goed. Scopus heeft een aparte sectie Environmental Science in de vakgebiedsindeling, maar afhankelijk van het onderwerp zitten ook milieukundige benaderingen bij Earth & Planetary Sciences, Social Sciences, Biological en Agricultural Science en ook bij Energy en Materials Science. Het is dus zaak hier niet a priori een inperking op vakgebied te gebruiken. Voor de innovatiewetenschappen (in hetzelfde departement als milieuwetenschappen) is een geïntegreerde database als Scopus ideaal. Scopus is hier beter dan WoS door grotere opname van niet-Amerikaanse tijdschriften en door goede dekking van technologie (op basis van Compendex) en economie. De domeinen van de innovatiewetenschappen (gentechnologie, energie en materialen en RO&transport) worden alle goed door Scopus gedekt. Er is voor deze vakgebieden geen gerichte set gecontroleerd vocabulaire beschikbaar, maar men zal wel profiteren van het vrij grote aantal trefwoorden uit Geobase, Compendex en Biobase.

7.2.10 Natuur- en Sterrenkunde

De dekking van natuur- en sterrenkunde is met 75% van de Utrechtse digitale tijdschriften vrij goed. De backfill moet echter beter worden, zeker gezien het feit dat er een grote hoeveelheid Open Access materiaal in dit vakgebied is en gezien de gratis beschikbaarheid elders van databases als ADS (Astrophysical Data System) en ArXiv. Weliswaar is er vanuit de onderliggende database Compendex backfill, maar dat is voor een groot deel techniek (engineering) en betreft maar in beperkte mate fundamenteel wetenschappelijke tijdschriften (die wel zijn opgenomen in Inspec). Aantallen citaties naar door ons geselecteerde artikelen komen overeen met die in WoS, maar enkele zoekacties uitgevoerd voor dit onderzoek zijn niet helemaal overtuigend. Dit behoeft nadere beschouwing.

Er is weliswaar een Scopus-vakgebied Physics and Astronomy, maar Materials, Energy en Earth & Planetary Science zullen ook vaak relevant zijn.

7.2.11 Rechtsgeleerdheid

Weliswaar hebben beide bestanden in absolute zin een tamelijk slechte dekking (Scopus met 35% nog wel iets minder slecht dan WoS) voor dit vakgebied, maar de onderlinge verhouding is toch wel interessant. Het is namelijk het enige vakgebied dat sterke discrepantie vertoont tussen dekking op basis van tijdschrifttitels en citatiegegevens (Uit tijdschrifttitels: Scopus/WoS= 1,6 en uit citaties: Scopus/WoS= 0,6 (en WoS/Scopus= 1,7).

Rechtsgeleerdheid wordt dus niet expliciet door Scopus gedekt maar er wordt ook niet als bij de humaniora over gezegd dat Scopus dat voorlopig ook niet ambieert. Waarschijnlijk speelt de bijzondere geardeerdheid van het materiaal (zeker voorzover gekoppeld aan de juridische praktijk binnen specifieke nationale contexten) een rol. Scopus kent weliswaar volgens de bronnenlijst 187 tijdschriften met de term law in de titel, maar de bulk daarvan is specifiek gericht op de VS. Wat betreft Scopus-vakgebieden zijn tijdschriften op het gebied van recht doorgaans geschaard onder social science.

7.2.12 Scheikunde

De scheikunde heeft een vrij goede dekking in Scopus: 75% van ons digitale scheikunde-bezit. Overige dekkingsindicatoren zijn ook redelijk goed tot goed, maar de backfill behoeft verbetering. Net als voor natuurkunde geldt hier dat de scheikundige titels die via compendex vulling leveren voor de periode vóór 1996 vooral toegepaste scheikunde betreffen en niet de vele meer fundamenteel wetenschappelijke titels.

Uiteraard kan Scopus niet tippen aan de dekking en functionaliteit van Chemical Abstracts of SciFinder Scholar, maar Scopus ondersteunt wel zoeken op CASREG-nummers en linking naar records in Crossfire Beilstein om reacties en molecuulstructuren te bekijken.

Voor scheikunde zijn er twee Scopus-vakgebieden: Chemistry (fundamenteel wetenschappelijk) en Chemical engineering (chemische technologie). Vaak zal echter ook het segment Materials science relevant zijn.

7.2.13 Sociale Geografie en Planologie

Voor sociale geografie en planologie heeft Scopus een goede dekking, zeker in vergelijking met WoS. Deze dekking gaat terug op Geobase. Ook de backfill is daardoor goed (tot 1980). Ook is hierdoor de beschikbaarheid van trefwoorden vrij groot. Ook geografische aanduidingen uit Geobase zijn meegenomen in Scopus.

Een probleem is de vakgebiedsindeling. Aanvankelijk zaten alle geowetenschappelijke tijdschriften onder Earth & environmental science, maar sinds de Spring 2006 release valt fysische geografie onder Earth & planetary science en sociale geografie en planologie onder Social science. Dat is op zich beter, maar veel van de voor SG&PL relevante tijdschriften zijn

toch alleen(ten onrechte) bij Earth & planetary science ingedeeld. Verder is en de rubriek Social science te breed om bij de meeste zoekacties van nut te zijn. Het is dus zelfs niet verstandig vooraf al een van de vier vakgebiedsaggregaten aan te vinken. Tenslotte: uit dit onderzoek blijkt dat Scopus voor dit vakgebied significant meer citaties naar publicaties in dit vakgebied kent dan WoS.

7.2.14 Sociale Wetenschappen

Dekking voor sociale wetenschappen is moeilijk als geheel te duiden. Interpretatie van de scores in tabel 7.1 is dan ook lastig. De dekking van tijdschrifttitels is overall met 54% redelijk, beter dan de 41% van WoS. Uit zoekacties op de hoofdtermen van de vakgebieden blijkt echter een minder goede dekking dan in WoS voor de maatschappijwetenschappen (tabel 7.2). Dit geldt zowel voor recentere als voor oudere jaren. Voor de psychologie is de dekking van oudere jaren bij Scopus beter en die van recente jaren bij WoS beter. Psychologie in Scopus profiteert waarschijnlijk van opname van psychologische tijdschriften in Scopus via EMBase.

Tabel 7.2 Zoekresultaat voor algemene termen uit de sociale wetenschappen als titelwoorden in Scopus en WoS, totaal en 1996-2005.

	alle jaren			1996-2005		
	Scopus	WoS	Scopus als % van WoS	Scopus	WoS	Scopus als % van WoS
anthropol*	3518	6051	58	1594	5050	32
sociolog*	4867	11570	42	2566	5895	44
psycholog*	44752	45402	99	19706	25336	78

Het beeld is wel weer anders op meer specifieke zoekacties binnen de vakgebieden (tabel 4.5). Daarbij scoort ook antropologie in Scopus goed, maar de dekking van sociologie roept ook daar weer vragen op.

Scopus mist de uitgebreide mogelijkheden en indexterminen van PsycInfo (nog afgezien van de dekking), maar heeft op citatiegebied toch iets meer functionaliteit. Voor sociale wetenschappen heeft Scopus twee vakgebiedsegmenten: Social Science (waaronder ook sociale geografie en recht) voor de maatschappijwetenschappen en Psychology voor de gedragswetenschappen. Niet is uitgezocht in welke Scopus-vakgebieden disciplines als pedagogiek en onderwijskunde onder zijn gebracht.

De situatie bij sociale wetenschappen behoeft nadere aandacht.

7.2.15 Wijsbegeerte

Van de humaniora is wijsbegeerte het vakgebied met de beste dekking, maar in absolute zin stelt het weinig voor. Er is enige dekking op het terrein van wetenschapsfilosofie, kunstmatige intelligente en wijsgerige antropologie. Op die terreinen zijn ook wel wat tijdschriften opgenomen.

7.2.16 Wiskunde en Informatica

De dekking voor wiskunde- en informatica in Scopus is niet meer dan redelijk, maar wel beter dan die van WoS op dit terrein. Hierbij moet wel worden gezegd dat de pure dekking voor het vakgebied waarschijnlijk iets groter is als gecorrigeerd wordt voor het vrij grote aantal tijdschriften op het gebied van bibliotheekwetenschap dat is opgenomen in dit UBU-vakgebied.

De Backfill voor dit vakgebied is in Scopus zeer beperkt: 10%. Voor informatie over ouder materiaal zal men vooralsnog Zentralblatt MATH moeten gebruiken, hoewel men daar

natuurlijk de citatiegegevens mist. Scopus is flink beter dan WoS wat betreft citaties naar wiskundige artikelen. Deze vakgebieden corresponderen met 3 Scopus-vakgebieden: Mathematics, Computer sciences en Decision sciences.

Literatuur

- Bakkalbassi, N., K. Bauer, J. Glover & L. Wang (2006) Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical Digital Libraries* 3,7. <http://www.bio-diglib.com/content/3/1/7>
- Deis, L.F. & D. Goodman (2006) Update on Scopus. *The Charleston Advisor* 7,3. <http://www.charlestonco.com/comp.cfm?id=55>
- Goodman, D. & L.F. Deis (2005) Web of Science (2004 version) and Scopus. *The Charleston Advisor* 6,3. <http://www.charlestonco.com/comp.cfm?id=43>
- Jascó, P. (2004) Scopus [online]. Péter's digital reference shelf, september 2004. <http://www.galegroup.com/servlet/HTMLFileServlet?imprint=9999®ion=7&fileName=reference/archive/200409/scopus.html>
- Jascó, P. (2005) As we may search - comparison of major features of the Web of Science, Scopus and Google Scholar citation-based and citation-enhanced databases. *Current Science* 89, pp. 1537-1547. <http://www.ias.ac.in/currsci/nov102005/1537.pdf>
- Jascó, P. (2006) Scopus revisited [online]. Péter's digital reference shelf, june 2006. <http://reviews.gale.com/index.php/digital-reference-shelf/2006/06/scopus-revisited/>
- Neuhaus, Chr., E. Neuhaus, A Asher & C. Wrede (2006) The depth and breadth of Google Scholar: an empirical study. *Portal: libraries and the Academy* 6, pp. 127-141. http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v006/6.2neuhaus.pdf
- Pipp, E (2006) Vergleich der von Scopus bzw. Web of Sciences erfassen Zeitschriften. *Online Mitteilungen* 85, pp. 3-17. <http://www.univie.ac.at/voeb/php/downloads/om85.pdf>
- Schneider, K. (2006) Scopus - Web of Science: Versuch einer Bewertung aus pharmakognostischer Sicht. *Online Mitteilungen* 85, pp. 21-24. <http://www.univie.ac.at/voeb/php/downloads/om85.pdf>
- Qin, J. (2000) Semantic similarities between a keyword database and a controlled vocabulary database: An investigation in the antibiotic resistance literature. *Journal of the American Society for Information Science* 51, pp. 166-180. <http://www3.interscience.wiley.com/cgi-bin/fulltext/69501138/PDFSTART>

In bijlage 6 staat overige, niet aangehaalde literatuur over Scopus

Bijlage VI: Niet-aangehaalde literatuur over Scopus

- Burnham, J.F. (2006) Scopus database: a review. Biomedical Digital Libraries 3,1. <http://www.bio-diglib.com/content/3/1/1>
- Dess, H.M. Database reviews and reports - Scopus. Issues in Science and Technology Librarianship, winter 2006. <http://www.istl.org/06-winter/databases4.html>
- Fingerman, S. (2005) Scopus: profusion and confusion. Online 29,2, pp.36-38. <http://www.infoday.com/Online/mar05/index.shtml>
- Goraiz, J. (2006) Web of Science versus Scopus oder das aktuelle Dilemma der Bibliotheken. Online Mitteilungen 85, pp. 25-30. <http://www.univie.ac.at/voeb/php/downloads/om85.pdf>
- Grupo SCImago (2006) Análisis de la cobertura de la base de datos Scopus. El profesional de la información 15, 2, pp.144-145. <http://www.ugr.es/~benjamin/EPI-Scopus.pdf>
- Kaemper, B-Chr. (2006) A Reader`s Reflection about Scopus: Letter from Bernd-Christopher Kaemper. The Charleston Advisor 7,4. <http://www.charlestonco.com/features.cfm?id=200&type=me>
- LaGuardia, C. (2005) ISI Web of Science / Scopus. Library Journal 130,1, pp.40-42. <http://www.libraryjournal.com/article/CA491154.html>
- Roth, D.L. (2005) The emergence of competitors to the Science Citation Index and the Web of Science. Current Science 89,9, pp. 1531-1536. <http://www.ias.ac.in/currsci/nov102005/1531.pdf>
- Wildner, B. (2006) Web of Science - Scopus: auf der Suche nach Zitierungen. Online Mitteilungen 85, pp. 18-20. <http://www.univie.ac.at/voeb/php/downloads/om85.pdf>