



Zitate, Downloads und Readership - eine Analyse auf Zeitschriftenebene

Dr. Peter Kraker

Bibliometrie 2014, 25.9.2014

Danksagung

Dieser Vortrag basiert auf

Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2014). Comparison of downloads, citations and readership data for two information systems journals. *Scientometrics*. doi:10.1007/s11192-014-1365-9

Schlögl, C., Gorraiz, J., Gumpendorfer, C., Jack, K., & Kraker, P. (2014). A comparison of citations, downloads and readership data for an information systems journal. *Research Trends*, (37).



Viele Studien haben Download- und Zitationsdaten verglichen (Kurtz et al. 2005, Bollen & Van De Sompel 2008, O’Leary 2008, Schlögl & Gorraiz 2011)

Online Referenzmanagement-Systeme sind als Altmetrics-Quelle hoch im Kurs

Viele Studien haben Zitations- und Readership-Daten miteinander verglichen (Bar-Ilan 2012, Li et al. 2012, Kraker et al. 2012, Mohammadi and Thelwall 2013)

In unseren Studien haben wir Zitations-, Download- und Readership-Daten für zwei Zeitschriften aus dem Bereich Information Systems verglichen

Sind die meist zitierten Artikel diejenigen, die am häufigsten downgeloadet/am häufigsten zu Mendeley hinzugefügt werden?

Wie verhalten sich Zitationen, Downloads und Readership über die Zeit und gibt es hier Unterschiede?

Gibt es andere Charakteristiken, in denen sich diese drei Indikatoren unterscheiden?

Daten

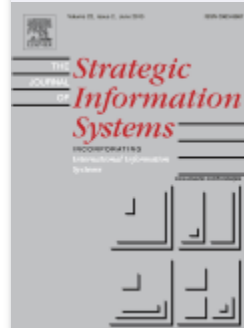
Zeitschriften

- The Journal of Strategic Information Systems (JoSIS)
- Information and Management (I&M)

**Analyseperiode: 2002-2011; n=968 Dokumente
(JoSIS: 321, I&M: 647)**

Datenquellen:

- ScienceDirect (SD): Downloaddaten, monatlich (PDF & HTML)
- Scopus: Zitationsdaten, monatlich
- Mendeley: Readership-Daten, monatlich (full length articles)



Vorverarbeitung

- Matching zwischen ScienceDirect and Scopus
 - Keine eindeutige ID für ScienceDirect und Scopus/unterschiedliche Dokumenttypen
 - Matching über Titel, Zeitschrift, Jahrgang/Ausgabe, Seiten
- Matching zwischen Scopus and Mendeley über den Titel (Levenshtein ratio 1/15.83); 69 nicht gefunden (5+64 \approx 7%)

Deskriptive Statistik

- Dokumenttyp, Publikationsdatum, Obsoleszenz, Nutzer

Korrelationsanalyse

- Downloads vs. Zitationen, Readership vs. Zitationen, Downloads vs. Readership

Ergebnisse

Downloads pro Dokumenttyp (JoSIS)

Document type	n	% docs	% downloads	Downloads per doc – relations
Announcement	5	1.6%	0.4%	5.9*y
Book review	4	1.2%	0.3%	5.5*y
Contents list	29	9.0%	0.4%	1.0*y
Editorial Board	29	9.0%	0.6%	1.5*y
Editorial	49	15.3%	3.3%	4.6*y
Erratum	1	0.3%	0.1%	5.7*y
Full length article	181	56.4%	94.1%	35.4*y
Index	12	3.7%	0.2%	1.3*y
Miscellaneous	9	2.8%	0.2%	1.8*y
Publishers note	2	0.6%	0.2%	7.0*y
	321	100%	100%	

Source: ScienceDirect; n=321

FLAs sind der am meisten downgeladete Dokumenttyp (94.1%)

Alle andere Dokumenttypen werden wesentlich seltener downgeloadet

Ergebnisse

Downloads pro Dokumenttyp (I&M)

Document type	n	% docs	% downloads	Downloads per doc – relations
Abstract only	1	0.2%	0.1%	14.8*y
Advertisement	4	0.6%	0.0%	1.0*y
Editorial Board	22	3.4%	0.2%	1.3*y
Editorial	1	0.2%	0.2%	34.6*y
Erratum	5	0.8%	0.2%	6.0*y
Full length article	581	89.8%	98.9%	29.0*y
Index	12	1.9%	0.1%	1.5*y
Miscellaneous	20	3.1%	0.2%	1.5*y
Short communication	1	0.2%	0.2%	35.7*y
	647	100%	100%	26.3*y

Source: ScienceDirect; n=647

FLAs sind der am meisten downgeladete Dokumenttyp (98.9%)

Es gibt aber zwei Ausreißer, die häufiger downgeloadet wurden als der ein durchschnittlicher FLA

Ausreißer 1: Peak, D.; Guynes, C.S.; Kroon, V. (2007)

Editorial: A profile of information systems research published in Information & Management. Information and Management, 44 (1), pp. 1-11.

Häufig downgeloadet und 88 mal zitiert

“[...] In a spirit of introspection, this article profiles research published in I&M: **we identified the most productive authors and universities associated with most research publications in I&M** during the past 13 years (1992-2005). Based on a more detailed analysis of publications during the past 7 years, we determined the subject areas most often investigated and the research methodologies most often employed. Finally, **we identified best practices by way of reporting the topics and methodologies used by the highly published authors.** [...]”

Ausreißer 2: Palvia, P.; Pinjani, P.; Sibley, E.H. (2005)

1. Full Article: Information Technology Alignment Planning—a case study. Information & Management, 42(4), pp. 619–633.

“The process is intended to aid in making the best possible use of IT resources in meeting the corporation's business objectives. [...] The model utilized **an intuitive color-coded alternative to statistical output that was readily accepted by management.**”

**Top 25 hottest articles ScienceDirect April–June 2005
Computer Science; 11 Zitationen**

2. Short Communication: Information Technology Alignment Planning—a case study. Information & Management, 42(5), pp. 635–649.

+20% bei Downloads

30 Zitationen

Ergebnisse

Verzögerung Printpublikation

Document types	J. of Strategic Information Systems		Information and Management	
	<i>n</i>	Online date—print publication date (mean days)	<i>n</i>	Online date—print publication date (mean days)
Abstract only			1	−260.0
Advertisement			4	−17.5
Announcement	5	−13.2		
Book Review	4	−40.5		
Content list	29	12.9		
Editorial Board	29	12.9	22	7.9
Editorial	49	9.0	1	−41.0
Erratum	1	−145.0	5	−31.2
Full length article	181	−49.8	581	−131.5
Index	12	−4.9	12	−4.3
Miscellaneous	9	32.9	20	−18.0
Short communication			1	−145.0
Publishers note	2	−13.0		
All	321	−24.9	647	−119.5

FLAs werden im Durchschnitt über 1.5 (JoSIS) bzw. über 4 Monate (I&M) vor dem Printdatum online publiziert

Ergebnisse

Downloads pro Publikationsjahr (JoSIS)

PY	n	DL-year											DL/ FLA
		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	all	
2002	13	1.0	2.3	1.7	1.3	1.2	1.4	2.4	2.8	2.8	2.7	19.6	7.4*y
2003	21	0.0	1.3	2.2	1.0	1.0	0.9	1.5	1.3	1.5	1.1	11.9	2.8*y
2004	17			1.7	2.6	2.1	2.2	2.4	2.7	2.9	2.3	18.9	5.5*y
2005	18				1.7	2.3	1.8	2.0	2.4	2.6	2.2	15.0	4.1*y
2006	14				0.2	2.4	2.1	1.8	2.1	2.0	2.0	12.5	4.4*y
2007	18					0.0	2.7	3.6	3.4	3.5	2.9	16.1	4.4*y
2008	16						0.0	2.9	3.5	3.0	2.4	11.8	3.6*y
2009	14								3.1	4.0	3.1	10.2	3.6*y
2010	21									3.9	4.4	8.3	2.0*y
2011	29									0.3	5.6	5.9	1.0*y
all	181	1.0	3.7	5.6	6.8	8.9	11.1	16.6	21.4	26.4	29.0	130.4	

Source: ScienceDirect; FLA only (n=181)

Download-Maximum in vielen Fällen 1 Jahr nach Publikation

Absolutes Download-Maximum für Artikel publiziert im Jahr 2011

Ergebnisse

Downloads pro Publikationsjahr (I&M)

PY	n	DL-year											DL/ FLA
		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	all	
2002	46	1.7	1.6	1.2	1.1	1.0	1.2	1.7	2.1	2.2	1.9	15.8	7.6*y
2003	73	0.5	3.1	2.2	1.5	1.3	1.4	1.9	2.2	2.2	1.9	18.3	5.6*y
2004	71		0.4	4.2	2.7	2.0	2.1	2.5	2.9	3.0	2.6	22.3	7.0*y
2005	61			0.6	3.6	2.1	1.7	2.0	2.5	2.4	2.1	17.1	6.2*y
2006	78				0.4	3.5	3.0	2.6	3.0	3.0	2.6	18.1	5.1*y
2007	48					0.0	2.6	2.2	2.0	2.1	1.7	10.7	4.9*y
2008	62						0.0	4.0	3.9	3.2	2.7	13.8	4.9*y
2009	56							0.0	3.8	3.1	2.4	9.3	3.7*y
2010	42								0.2	2.9	2.1	5.2	2.8*y
2011	44									0.0	2.0	2.0	1.0*y
all	581	2.2	5.1	8.3	9.2	9.9	12.0	16.9	22.7	24.1	22.0	132.5	

Source: ScienceDirect; FLA only (n=581)

Download-Maximum in den meisten Fällen im Jahr der Publikation

Absolutes Download-Maximum für Artikel publiziert im Jahr 2004

Ergebnisse

Zitationen pro Dokumenttyp

JoSIS	no. docs	% uncited	Cites	Cites per doc type
Article	151	15%	2563	14.8
Conference paper	13	69%	8	0.4
Editorial	33	79%	13	0.2
Review	18	6%	383	20.2
All	215	27%	2967	10.9

Source: Scopus; n=215

I&M	no. docs	% uncited	Cites	Cites per doc type
Article	549	9%	14431	26.3
Conference paper	1	0%	43	43.0
Editorial	5	80%	1	0.2
Review	39	8%	1661	42.6
All	594	10%	16136	27.2

Source: Scopus; n=594

**Unterschiedliche Dokumenttypen in Scopus und ScienceDirect
(FLA \approx Articles + Conference Papers + Reviews)**

27% bzw. 10% der Dokumente wurden nie zitiert (besonders Editorials, Conference Papers und neuere Artikel)

Ergebnisse

Zitationen pro Publikationsjahr (2002-2011)

Special Issue: "Trust in the Digital Economy"

Pub year	n	Citation year											cites per doc
		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	all	
2002	13	2	19	38	69	88	105	158	165	194	199	1037	79.8
2003	14		1	6	21	27	39	35	41	40	39	249	17.8
2004	17			0	15	40	56	74	78	88	107	458	26.9
2005	19				0	16	46	78	76	93	99	408	21.5
2006	14				1	2	14	31	31	53	49	181	12.9
2007	18						1	31	74	92	85	283	15.7
2008	15							3	30	69	83	185	12.3
2009	14								3	34	57	94	6.7
2010	18									5	40	45	2.5
2011	8										14	14	1.8
all	150	2	20	44	106	173	261	410	498	668	772	2954	

Special Issue mit Conference-Papers

Source: Scopus; Document types: articles, reviews, conference papers; only cited documents (n=150)

Nur wenige Dokumente werden im Publikationsjahr zitiert – Zitationsmaximum wird mehrere Jahre nach der Publikation erreicht

Unterschied zwischen Downloads: Maximum im Publikationsjahr oder ein Jahr später

Ergebnisse

Zitationen pro Publikationsjahr (I&M)

Pub year	n	Citation year											cites per doc
		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	all	
2002	46	3	32	35	38	40	40	37	40	40	38	2617	56.9
2003	72		12	43	59	58	63	67	61	62	66	2826	39.3
2004	72			6	54	64	70	70	65	71	69	3603	50.0
2005	62				14	47	51	55	58	60	56	2278	36.7
2006	77					14	52	61	69	75	69	1985	25.8
2007	54						11	45	49	53	54	1227	22.7
2008	61							20	49	55	60	988	16.2
2009	50								14	44	46	435	8.7
2010	33									12	32	125	3.8
2011	6										6	8	1.3
all	533	3	44	84	165	223	287	355	405	472	496	16092	

Source: Scopus; Document types: articles, reviews, conference papers; only cited documents (n=533)

Nur wenige Dokumente werden im Publikationsjahr zitiert – Zitationsmaximum wird mehrere Jahre nach der Publikation erreicht

Unterschied zwischen Downloads: Maximum im Publikationsjahr oder ein Jahr später

Ergebnisse

Readership pro Publikationsjahr

Publication year	Journal of Strategic Information System			Information and Management		
	<i>n</i>	Occurrences in user libraries	Occurrences per doc	<i>n</i>	Occurrences in user libraries	Occurrences per doc
2002	13	566	43.5	46	992	21.6
2003	21	344	16.4	73	1,168	16.0
2004	17	471	27.7	71	1,701	24.0
2005	18	371	20.6	61	1,071	17.6
2006	14	382	27.3	78	1,669	21.4
2007	18	580	32.2	48	909	18.9
2008	16	451	28.2	62	1,317	21.2
2009	14	416	29.7	56	911	16.3
2010	21	499	23.8	42	727	17.3
2011	29	537	18.5	44	484	11.0
All	181	4,617	25.5	581	10,949	18.8

Source: Mendeley, FLA only

Ergebnisse

Readership pro Publikationsjahr (JoSIS)

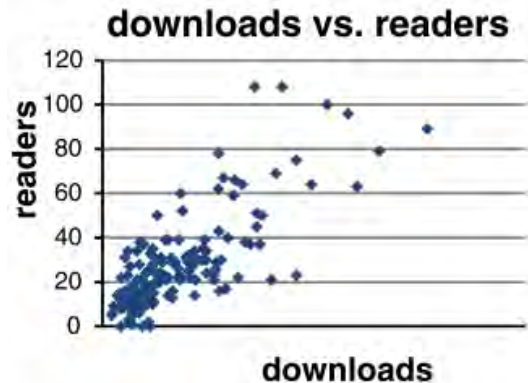
Pub year	n	Readership years					all	Readers per doc
		2008	2009	2010	2011	- July 2012		
2002	13	7	30	126	245	183	591	45.5
2003	21	1	29	58	108	145	341	17.1
2004	17	11	36	107	158	165	477	28.1
2005	18	2	31	79	141	151	404	23.8
2006	14	6	39	88	128	148	409	29.2
2007	18	4	45	129	222	209	609	35.8
2008	16	7	36	99	182	164	488	32.5
2009	14	0	27	111	127	150	415	29.6
2010	21	0	0	84	238	191	513	24.4
2011	29	0	0	4	208	282	494	17.6
all	181	38	273	885	1757	1852	4741	

Source: Mendeley; FLA only (n=181)

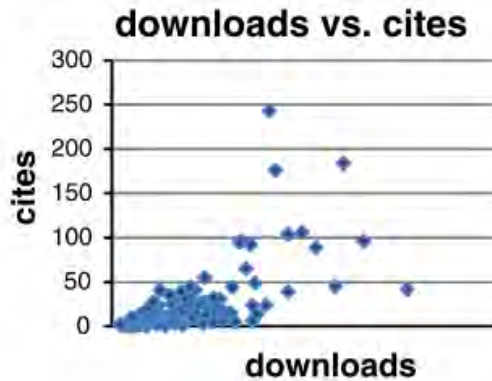
Das geringe Alter von Mendeley (gegr. 2008) und der starke Anstieg in der Nutzerbasis seitdem (jetzt: 2.8 mio) machen Obsoleszenzanalysen schwer – Gewichtung wäre notwendig

Ergebnisse

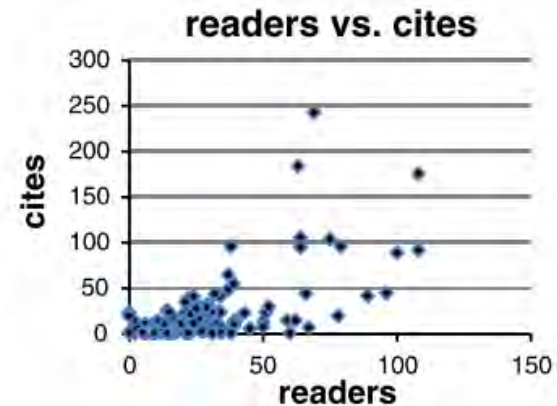
Downloads vs. readers vs. cites (only FLAs)



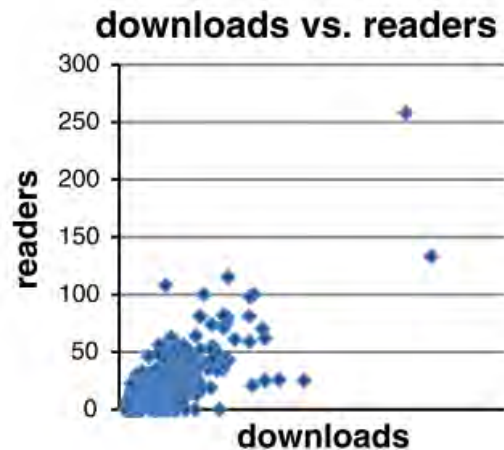
$r=0.73$, $n=150$



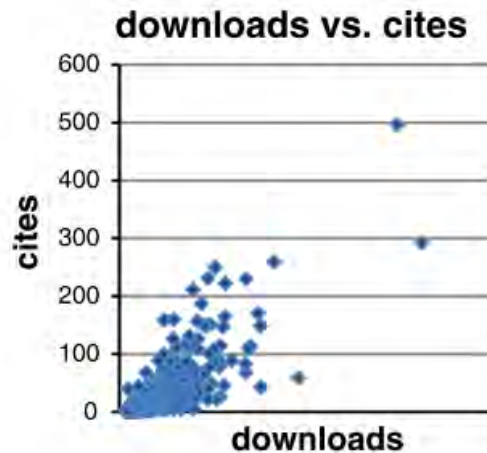
$r=0.77$, $n=150$



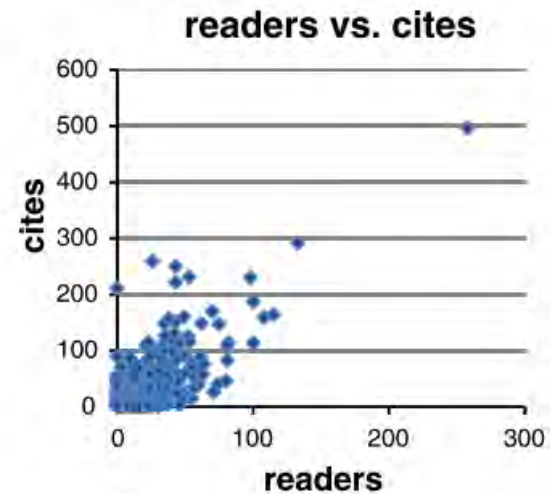
$r=0.51$, $n=150$



$r=0.66$, $n=528$



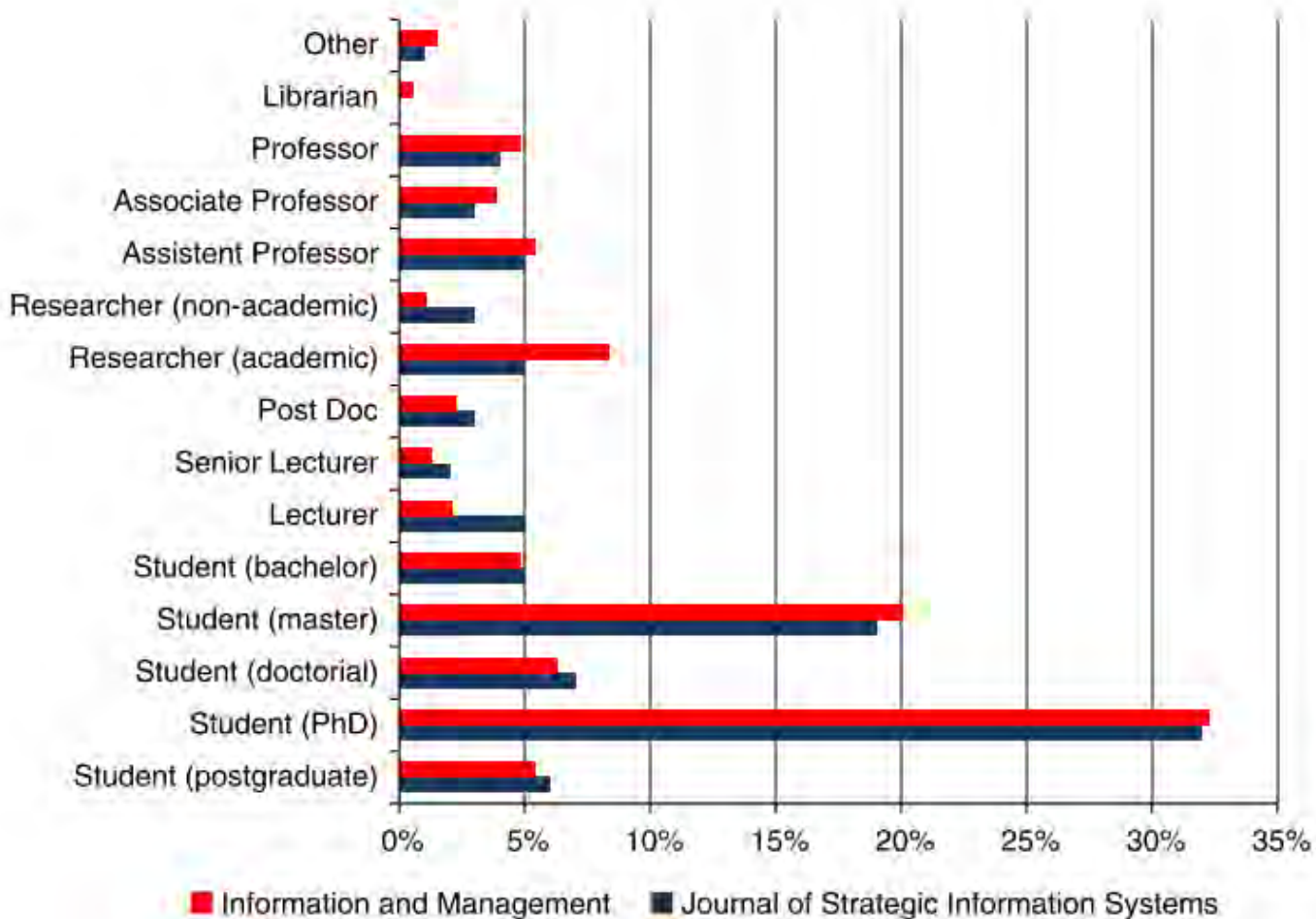
$r=0.76$, $n=528$



$r=0.59$, $n=528$

Ergebnisse

Readership nach Nutzergruppen (Mendeley)



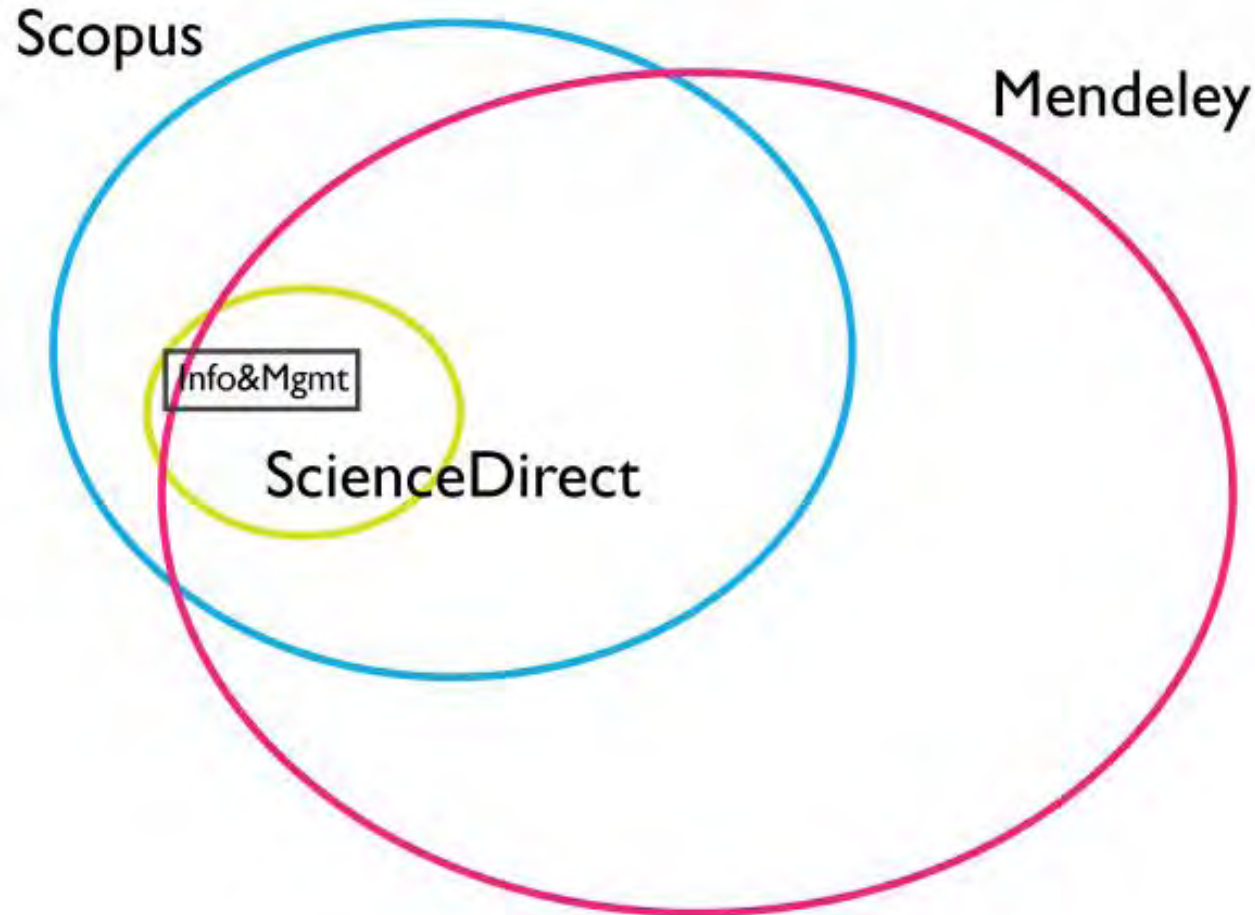
Source: Mendeley; doc type: FLA

2/3 der Readership kommt von Studierenden

Researchers + Post Docs + Profs \approx 1/4 der Leserschaft

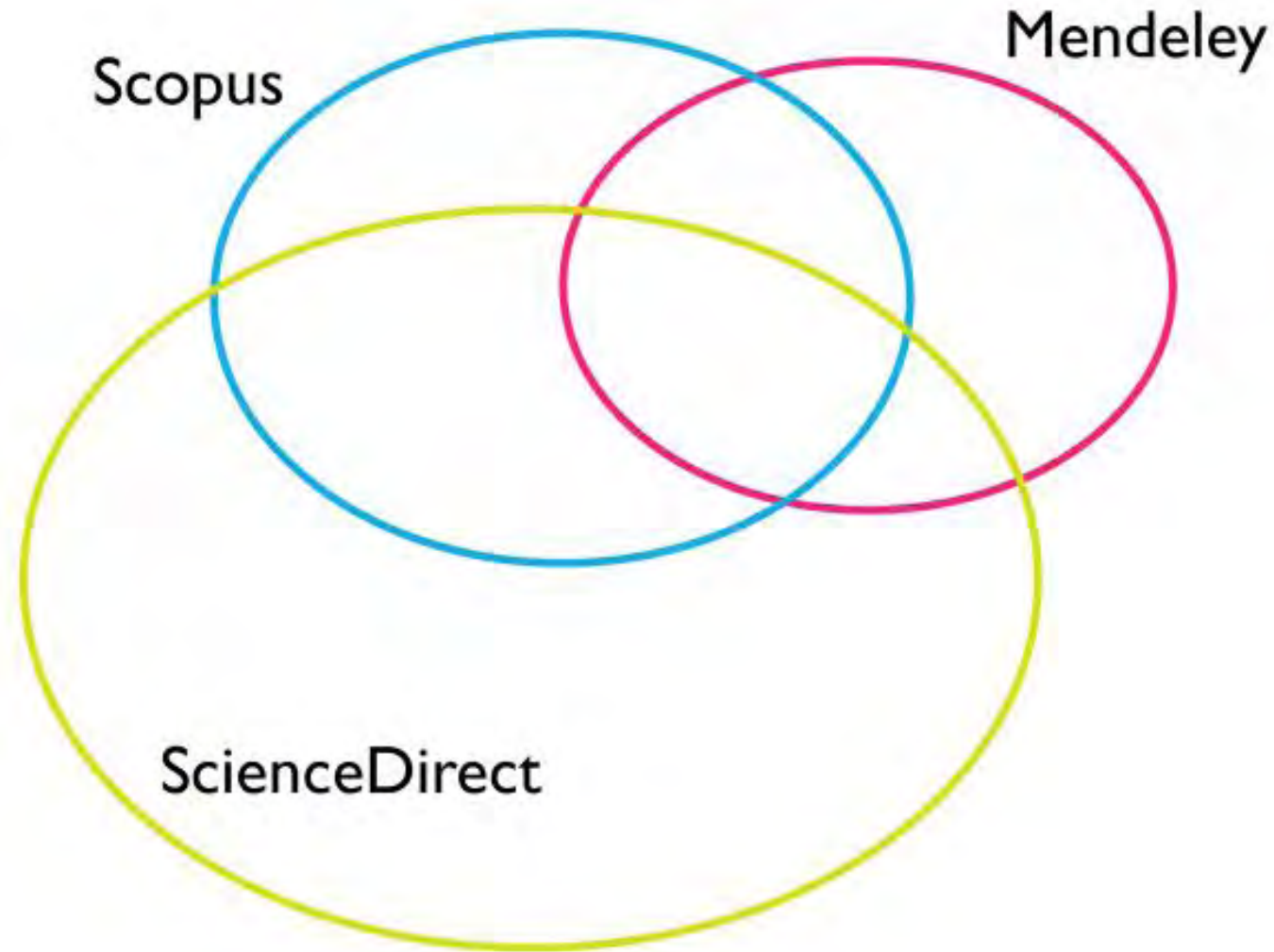
Diskussion

Abdeckung der Literatur durch Systeme



Diskussion

Unterschiedliche Nutzer-Grundgesamtheiten



Schlussfolgerungen

Vergleich unterschiedlicher Maße ist mitunter schwierig

Verschiedene Obsoleszenz-Charakteristiken von Downloads und Zitationen (Readership steht noch aus)

Mittlere bis hohe Korrelation zwischen Downloads und Zitationen

Mittlere Korrelation zwischen Readership und Zitationen

Erklärungsansätze

- Unterschiedliche Nutzergruppen
- Unterschiedliche Funktionen der einzelnen Maße

Um diese Unterschiede besser zu verstehen, benötigen wir offene Daten, die es uns erlauben, auf den Kontext einzugehen

Vielen Dank für die Aufmerksamkeit!



Dr. Peter Kraker

Scientific Project Development Science
2.0 and Open Science

pkraker@know-center.at

<http://twitter.com/PeterKraker>

<http://science20.wordpress.com>