MAX PLANCK
digital library

ScholarlyStats (Swets)
Erfahrungen und Bewertungen aus der Sicht der Max-Planck-Gesellschaft

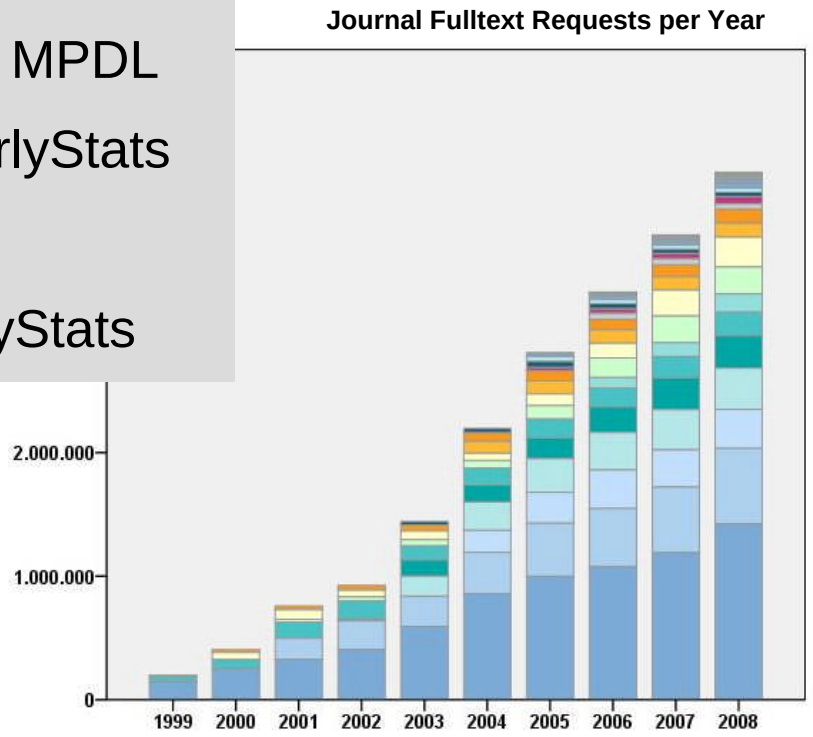
3. Konsortialtag des Friedrich-Althoff-Konsortiums e.V.
Nutzungsdaten – Raum für Kreativität?
Theorie und Praxis bei Anbietern und Bibliotheken
Berlin, 21. April 2009

Mag. Margit Palzenberger
Max Planck Digital Library
Scientific Information Provision

Einsatz von ScholarlyStats in der MPG

- 1999 Start zentrale Lizenzierungen
- 2003-2005 erhebliche Erweiterung
- 2005 Jan Gründungsbeschluß MPDL
- 2005 Okt Lizenzierung ScholarlyStats
- 2007 Jan Gründung MPDL
- 2008 Nov Kündigung ScholarlyStats

➔alle weiteren Ausführungen auf Basis des Status von Ende 2008 !!



Gründe für Lizenzierung / Abbestellung

- Lizenzierung
  - bei extrem geringer personeller Ausstattung
  - bei stark anwachsendem Portfolio
  - ➔ möglichst unaufwändige Sicherung der anfallenden Daten
  - ➔ schnelle "Grob-Sichtung" der Nutzung
  - ➔ Webinterface für verteilte Zugriffe

- Abbestellung
  - Konsolidierung des ERM und damit der notwendigen Infos
  - Harvesting kann nun über HiWis organisiert werden
  - ScholarlyStats verursacht mehr Aufwand als erwartet
  - ScholarlyStats bietet einige kritische Funktionen nicht
  - ➔ Ablöse durch Eigenentwicklung

Profil von ScholarlyStats

in der Zeit unserer Lizenzierung war ScholarlyStats für uns

ein

- Harvester für COUNTER-Daten
- Archiv für diese Daten
- Aggregator für diese Daten
- standardisiertes Web-Interface zum Download dieser Daten

kein

- Nutzungsdaten-Management-System
- Tool für adäquate statistische Analysen
- Reporting- und Visualisierungs-Tool

ScholarlyStats als Harvester

Publisher/Platform: **Elsevier ScienceDirect**

*Username for collection: xxxxxxxxx

Password for collection: xxxxxxxxx

*URL for collection: http://usagereports.elsevier.com

Publisher	Platform	MPS Collection Platform	COUNTER	Report Measurement	Journal Title	Print ISSN	Online ISSN	Jan 07	Feb 07	Mrz 07	Apr 07	YTD-07
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	Nature	"0028-0836	"1476-4687	16523	14654	16089	16776	64042
American Association For T	HighWire Press	American Associati	yes	Full-Text Requests	Science	"0036-8075	"1095-9203	14426	12426	13684	14296	54832
Aps	Scitation	Scitation	yes	Full-Text Requests	Physical Review Online Archive	no data	no data	6815	8114	9971	7942	32842
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Cell	"0092-8674	no data	6045	5969	5594	5045	22653
Aps	Scitation	Scitation	yes	Full-Text Requests	Physical Review Letters	"0031-9007	"1079-7114	3731	3763	4287	3706	15487
Wiley	Wiley Interscienc	Wiley InterScience	no	Full-Text Requests	Angewandte Chemie International Edition	"1433-7851	"1521-3773	3623	2765	3402	2779	12569
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Current Biology	"0960-9822	no data	3159	2718	3200	3104	12181
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Neuron	"0896-6273	no data	2830	2441	2740	2528	10539
Aps	Scitation	Scitation	yes	Full-Text Requests	Physical Review B	"1098-0121	"1550-235X	2141	2268	2660	2294	9363
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	News@Nature.Com	"0000-0000	"1744-7933	2421	2112	2760	1928	9221
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Journal Of Molecular Biology	"0022-2836	no data	2175	2128	2247	1844	8394
Aip	Scitation	Scitation	yes	Full-Text Requests	The Journal Of Chemical Physics	"0021-9606	"1089-7690	1815	1939	2233	2100	8087
Aip	Scitation	Scitation	yes	Full-Text Requests	Applied Physics Letters	"0003-6951	"1077-3118	1919	1709	2079	1651	7358
Oxford Journals	HighWire Press	Oxford Journals Onl	yes	Full-Text Requests	Nucleic Acids Research	"0305-1048	"1362-4962	1871	1891	1728	1385	6875
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Molecular Cell	"1097-2765	no data	1984	1664	1735	1323	6706
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	The Embo Journal	"0261-4189	"1460-2075	1477	1514	1724	1373	6088
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Tetrahedron Letters	"0040-4039	no data	1528	1561	1505	1307	5901
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	Nature Neuroscience	"1097-6256	"1546-1726	1571	1394	1448	1352	5765
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Febs Letters	"0014-5793	no data	1518	1225	1436	1278	5457
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	Nature Biotechnology	"1087-0156	"1546-1696	1516	1106	1396	1337	5355
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	Nature Cell Biology	"1465-7392	"1476-4679	1408	1279	1348	1179	5214
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	Nature Methods	"1548-7091	"1548-7105	1206	1277	1436	1240	5159
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	Nature Genetics	"1061-4036	"1546-1718	1362	1302	1257	1121	5042
Nature Publishing Group	nature.com	Nature Publishing G	yes	Full-Text Requests	Nature Reviews Molecular Cell Biology	"1471-0072	"1471-0080	1438	1122	1377	1089	5026
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Biochemical And Biophysical Research C	"0006-291X	no data	1258	1264	1293	1094	4909
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Surface Science	"0039-6028	no data	1340	1266	1210	1015	4831
Blackwell Publishing	Blackwell Synerg	Blackwell Synergy	yes	Full-Text Requests	The Plant Journal	"0960-7412	"1365-313X	1336	1178	1170	1096	4780
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Neuroimage	"1053-8119	no data	1240	1145	1128	1144	4657
Elsevier	Elsevier Science[	Elsevier ScienceDir	yes	Full-Text Requests	Developmental Cell	"1534-5807	no data	1266	1116	1138	1104	4624

ScholarlyStats als Web-Interface für Downloads

The latest ScholarlyStats reports for Max Planck Society, Apr-2007:

Consolidated Reports	Description	ZIP	CSV	XLS
Consolidated Journal Report 1	Full-Text Article Requests by Journal Title, Platform and Month	(272 KB)	(1 MB)	(7 MB)
Consolidated Database Report 1	Searches and Sessions by Month and Database	(2 KB)	(8 KB)	(119 KB)
Consolidated Database Report 2	Turnaways by Month and Database	(0.8 KB)	(2 KB)	(93 KB)
Consolidated Database Report 3	Searches and Sessions by Month and Service	(0.71 KB)	(1 KB)	(80 KB)

Dashboard Reports	Description	ZIP	CSV	XLS
Dashboard Report 1	Number of Titles by Platform	(0.45 KB)	(0.42 KB)	(153 KB)
Dashboard Report 2	Number of Full Text Article Requests by Platform	(0.67 KB)	(0.8 KB)	(155 KB)
Dashboard Report 3	Average Use By Platform	(0.66 KB)	(0.77 KB)	(154 KB)
Dashboard Report 4	Top Use Journals by Platform	(3 KB)	(7 KB)	(192 KB)
Dashboard Report 5	Low Usage Journals (including zeros)	(151 KB)	(611 KB)	(3 MB)
Dashboard Report 6	Zero Usage Journals Report	(101 KB)	(374 KB)	(2 MB)
Dashboard Report 7	Low Usage Journals (excluding zeros)	(49 KB)	(185 KB)	(1 MB)
Dashboard Report 8	Top 50 journals across platforms	(1 KB)	(4 KB)	(162 KB)
Dashboard Report 9	Proportional Usage of Journal Titles	(148 KB)	(695 KB)	(3 MB)
Zip File of all Reports	Zip file containing all of the above reports	(4 MB)		

Problemfelder I

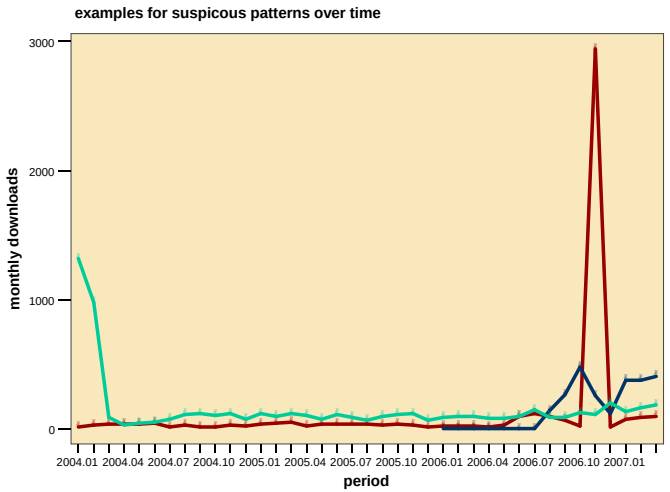
- nur ein Account pro Plattform
  (keine Instituts-Sichten, keine Teilpakete)
- keine Möglichkeit "alte" Daten hochzuladen
- keine Zusammenführung / Auswertung über Jahre hinweg
- keine Korrektur/Anreicherung durch Anwender möglich
- keine Korrektur/Anreicherung/Normierung durch ScholarlyStats
  - Zusammenführung von gewanderten/geänderten Titeln
  - Ergänzung von fehlenden ISSNs
  - einheitliche Behandlung von Backfile-Titeln
  - Behandlung von Ausreissern
  - Behandlung von Sonderzeichen

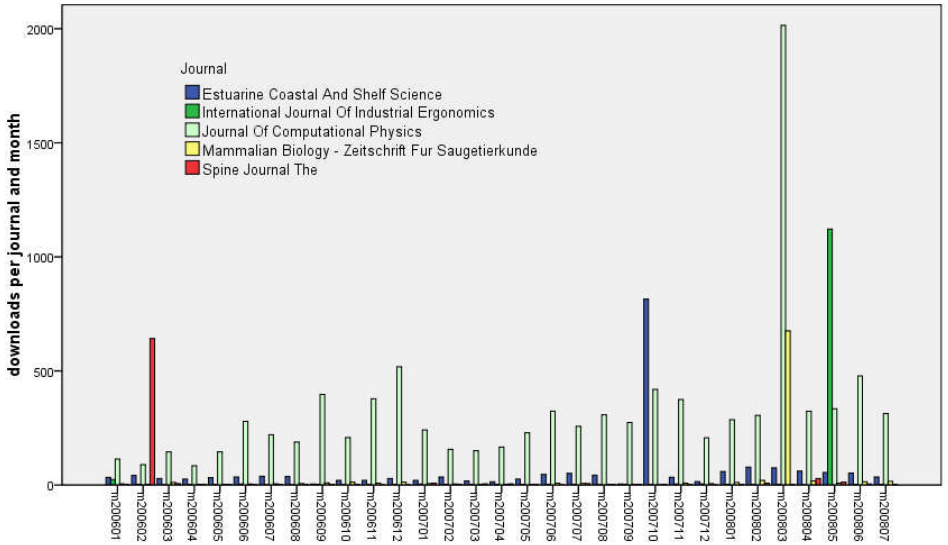
Problemfelder II

- nur einfachste "statistische" Kennwerte
  - Rankings
  - arithmetische Mittelwerte

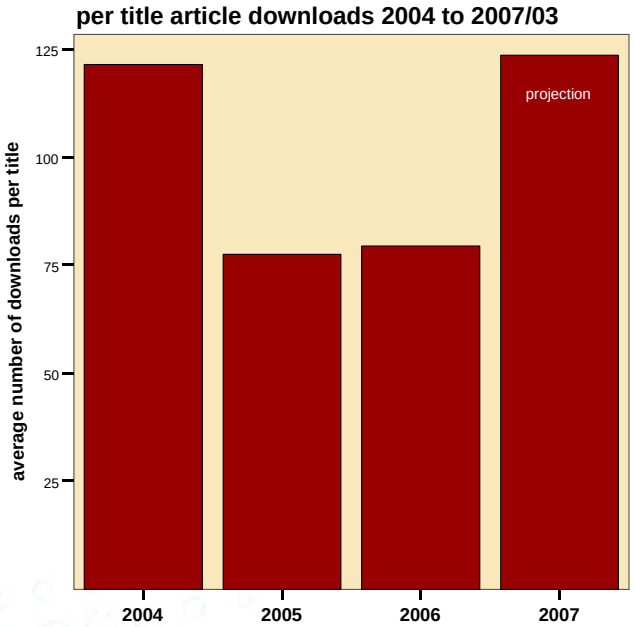
- diese können sehr irreführend sein
  - immer wieder einzelne Extremwerte ("Ausreisser") durch Mißbrauch, Log-Fehler, Fehler beim Datentransfer, …
  - auch im Normalfall extrem schiefe Häufigkeits-Verteilungen
  - Verlagsspezifika bei der Datenerfassung (Backfiles, …)

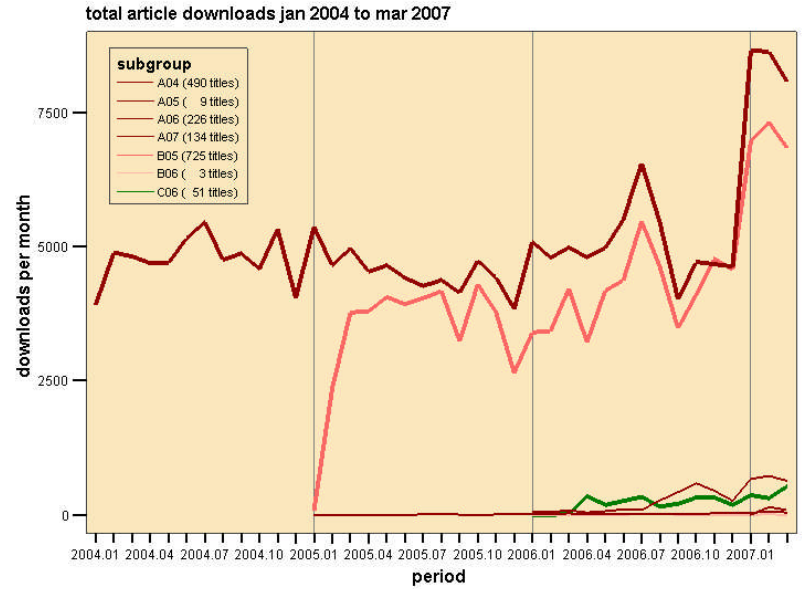
Ausreisser und andere Störenfriede …

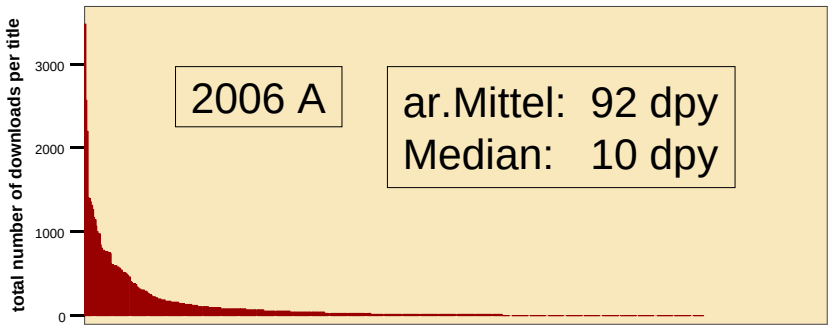




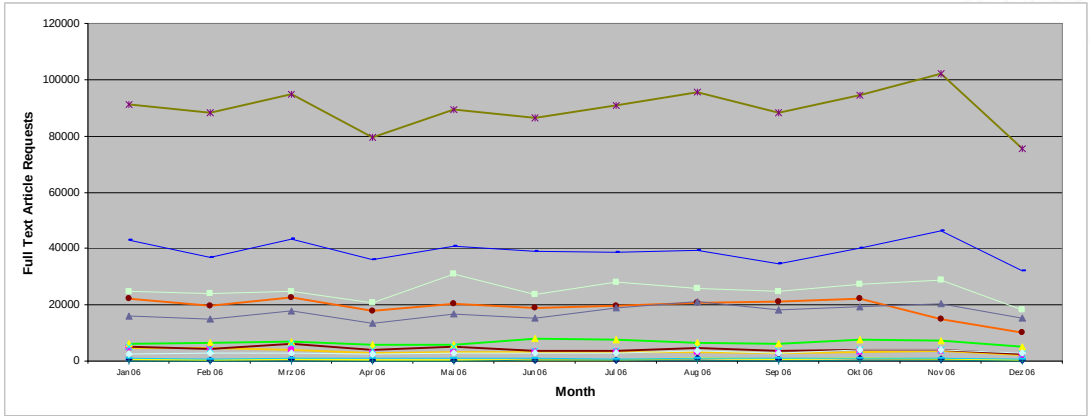
die Mitte finden …





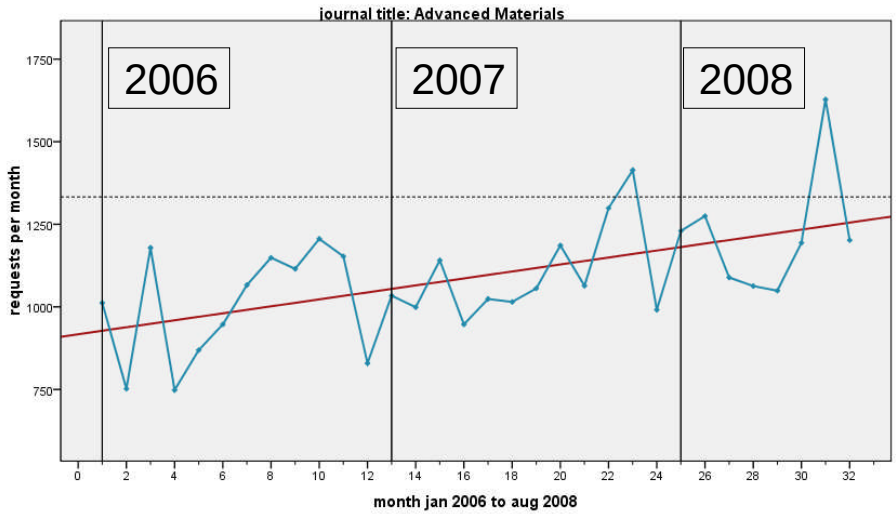


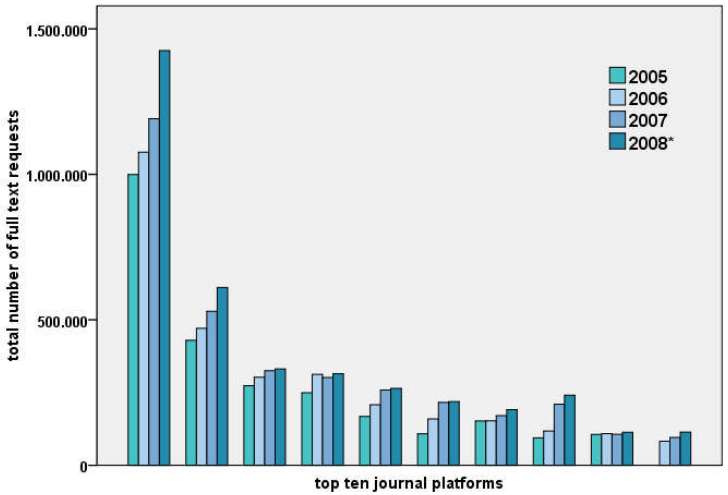
ein Jahr ist kein Jahr …



links:
ScholarlyStats Excel-Graphik
Plattformen pro Monat

unten:
Bsp für SPSS Graphiken zur zeitlichen Dynamik





… und ganz generell …

- Zusammenführung von Nutzungsdaten zahlreicher wissenschaftlicher Institutionen bei einem Anbieter
  - ➔ umfassender Markt-Einblick
  - ➔ eigentlicher Geschäftszweck ??

- mögliches gemeinsames Ziel: OpenSource SUSHI-Client mit den wesentlichen Funktionen von ScholarlyStats (Harvesting, Aggregation)