

Status Report - CSA

(October 2005)

Institute (Publisher):

CSA
7200 Wisconsin Avenue, Bethesda, MD 20814, USA
1-800-843-7751, 1-301-961-6740 (fax)
<http://www.csa.com>

Board Member & Staff:

President -- Matt Dunie
Chief Executive Officer – Jim McGinty

Main ASFA Contacts at CSA:

Craig Emerson (VP Editorial: CEmerson@csa.com)
Vicki Soto (Aquatic/Meteorological Sciences Supervisor: VSoto@csa.com)

Trust Fund Royalties

CSA is very pleased to report that ASFA Trust Fund revenues for the 2004 publication year increased by greater than US\$13,000 relative to 2003. A cheque for US\$212,998.00 payable to FAO for the ASFA Trust Fund was sent to FAO in April of this year, representing royalties as prescribed in the Publishing Agreement.

Input Statistics (Annex 1):

CSA contributed 22,773 records in 2004. Twenty-four ASFA Partners, including CSA, contributed fewer records than in the previous year. For our part, CSA has hired additional staff in 2005 and we are confident that our input will surpass previous years. Five partners increased their input significantly, most notably the Partner from China. (Five partners did not contribute any records during 2004.) The total of all partners' contributions was 38,225, a decrease of approximately 4% from last year.

Receipt of Partners' Input:

We encourage partners to input their ASFA records in Web-ISIS format. As always, please make sure that your email notification includes which format was used to create the records; absence of this information will delay input into the database.

To ensure timeliness, we would like to receive smaller, more frequent batches of records from partners who usually send their annual input in one or two large batches. We can also detect, correct and generally mitigate errors more efficiently in this manner.

Partners should make every effort to review their files before transmittal to CSA. It is mandatory that Field 91 be marked C for Complete and not T for Temporary when the record is exported. Records submitted with a T in this field are considered incomplete by the conversion program and are not loaded.

In most cases, we're pleased to report few problems with partner record loads, but we would like all partners to view their input statistics (www.csa.com/asfa/summary.htm; username/password asfaboard/123asfa) and report any problems to us as quickly as possible. An e-mail message to Vicki Soto (VSoto@csa.com) with a CC: to Craig Emerson (CEmerson@csa.com) and Hemant Jain (HJain@csa.com) should accompany all input.

If you are requesting corrections to abstract records, please remember to send an e-mail to Vsoto@csa.com with "record correction" in the subject line of the e-mail. Please include the following information:

- All pertinent fields: TR, TI, AN and UD (if available)
- Description of error
- Correction that you would like to have made

CSA's Monitoring List:

CSA currently monitors 3,172 journals for ASFA, an increase of approximately 78 titles since last reported. Many of these titles are covered selectively, but several will have significant coverage. For example:

- Ocean Science
- Ocean Science Journal
- Marine Biology Research
- PLoS Biology

The full 'unofficial' list, including journals monitored by all partners, can be found at http://www.csa.com/ids70/serials_source_list.php?db=aquclust-set-c. It includes over 5,300 titles. The full *official* monitoring list can be found at FAO's web site: <http://www.fao.org/fi/asfa/asfa.asp#ASER>.

Who is Using the ASFA Database? – CSA Illumina Usage Reports:

Note that with the 2005 launch of the new Web platform, the Internet Database Service (IDS) was officially renamed “CSA Illumina”.

The number of CSA Illumina searches in the ASFA database increased to 2,199,832. Searches originated from over 71 countries, with the USA responsible for 41% of the usage.

The largest single institutional ASFA user was Fisheries & Oceans, Canada, followed closely by National Oceanic & Atmospheric Administration, USA and University of Washington, USA. After removing USA searches, the next ‘top-5’ country usage is as follows:

Country	% of Total Use
Canada	16%
UK	14%
Australia	10%
France	6%
Mexico	5%

ASFA Partner Usage on CSA Illumina:

The following shows the ‘top-10’ CSA Illumina usage for the ASFA partnership, as a percentage of the ASFA Input Centre total, excluding NOAA, USA and Fisheries & Oceans, Canada (mentioned above).

Organization	Query
IFREMER SDP Brest Service	33%
Inst Marine Res	11%
Natl Ctr Marine Res	9%
Plymouth Marine Lab Oratory	7%
Inst Marine Res – Sweden	5%
Inst Nacl Invest y Desar Pesqu	5%
Food & Agriculture Org-UN	4%
KORDI	4%
Univ South Pacific	3%
Flanders Marine Inst	2%

Sales Information

In 2004, over 400 customers had subscriptions to at least one of the components of the ASFA database, with many more subscribing to clustered products which include ASFA files as a subset. Sales for ASFA-only products are available in Annex 2.

Sales & Marketing:

See Annex 3 and 4.

New Developments

Enhancements to Online ASFA: “CSA Illumina”

- Indexing terms are more prominent
- Eliminated frames (Advantages include easier/faster development time and better performance)
- Improved multi-language interface management and presentation.
- Advanced Search: End-users are no longer limited by the forced ‘AND’ between the first and second search rows. They will have a drop-down box with ‘AND’, ‘OR’, and ‘NOT’.
- First few lines of abstracts appear on results page
- Search results displayed by publication type tabs
- Peer reviewed journals are now identified
- Number of search results displayed per page can be customized
- Cross-database search results can be viewed in a combined list
- Search box is shown on the bottom of the page; easy to refine search
- Search rows can be added or removed in advanced search
- Improved access to Alerts set-up options
- Command-line search moved to Professional search
- Users can easily combine searches or submit a new search from the Search History page
- More full-text linking options – new publishers added:
 - American Geophysical Union
 - American Institute of Physics
 - American Mathematical Society
 - American Society for Microbiology
 - American Society for Nutritional Sciences
 - ANSInet
 - BMJ Publishing Group
 - Cambridge University Press
 - EBSCOHost
 - -Alt-HealthWatch
 - -Communication & Mass Media Complete
 - -Health Source: Consumer Edition
 - -Health Source: Nursing/Academic Edition
 - -Military & Government Collection
 - -Nursing & Allied Health Collection: Basic
 - -Nursing & Allied Health Collection: Comprehensive

- -Nursing & Allied Health Collection: Expanded
- -Professional Development Collection
- EDP Sciences
- Gale Group
- Haworth Press Inc
- Highwire Press, 1 new package
 - -Free Access Journals
- Hogrefe & Huber Publishers
- Institute of Physics
- NRC Research Press
- Project Muse, 1 new package
 - -Basic Undergraduate Collection
- ProQuest, 5 new packages:
 - -Academic Research Periodicals
 - -Canadian Serials
 - -ProQuest Medical Library
 - -ProQuest Nursing Journals
 - -ProQuest Psychology Journals
- Public Library of Science
- Society of Exploration Geophysicists
- SIAM - Society for Industrial and Applied Mathematics Taylor and Francis Group
University of Chicago Press Wildlife Disease Association

Cited References

In May 2005, the cited references appearing at the end of journal articles were targeted for capture. To date, the cited references for approximately 25,000 articles have been input into our production system. It is expected that cited references will be available to ASFA users in 2006.

Recent References

To narrow the time gap between the date when an article is published and the date when ASFA users are able to see that article abstracted and indexed in CSA Illumina, CSA purchases citations from the British Library. Although they are citation-only, "Recent References" from hundreds of core titles are added to CSA Illumina within a few days of the fulltext being published.

In Fall 2005, CSA will improve upon this Recent Reference capability by generating the references directly from CSA's production system. Whether or not an abstract record is indexed, the record will be available for searching. These Recent References will undergo quality checks and will include an "Indexing in process" tag to distinguish them from a completed record. Once indexed, the completed abstract record will replace the Recent Reference. Several thousand records will be available to ASFA users, weeks or months before they would normally see them.

Datasets Database

In early 2006, CSA editors will be identifying Web sites and other source of information containing datasets, the raw or minimally-processed data collected by researchers for analysis. More and more of these datasets are available in institutional repositories, or on individual Web sites, as well as published as appendices in the journal literature. In the past several months, A&I customers have requested that we include links to this information; such links will greatly enhance the value of our A&I offerings.

ASFA on CSA's CD:

Although the production of CDs are commercially untenable, CSA has made a commitment to make the ASFA database available in a CD product. (The ASFA database remains the sole CD product produced at CSA.) Although the CD is usable 'as is', there have been several suggestions by ASFA partners as to how the CD could be improved. Unfortunately, the software architecture used in the present version does not allow some of the more pertinent suggestions to be addressed (e.g. the inability to search both CD's simultaneously).

To resolve the above concerns, our Chief Technology Officer is investigating new software packages which may suit the needs of the ASFA CD users in a more efficient way. Additionally, we have approached NISC with the objective of creating an agreement whereby NISC would supply an ASFA-only CD to the ASFA community. This coincides with the contract renewal cycle between CSA and NISC. CSA hopes that the discussions continue to be positive and will result in a mutually beneficial arrangement. The current arrangement ends in December 2005, and it is likely that we will have an answer for the Board in late November.

Competitive Landscape

A&I products are becoming increasingly commoditized – undifferentiated – to the extent that unless they contain unique offerings, many A&I databases will not longer be viable, and users will be faced with very few choices of A&I publishers.

ISI's Web of Science (Thomson publishing group) and Scopus (Elsevier) are the dominant A&I publishers of Life Science records, and their recent marketing strategies have collateral impacts on smaller A&I publishers. For example, both WofS and Scopus offer their very expensive products at a remarkably low price, at least temporarily in an effort to dislodge the competition. Although the direct targets are each other, smaller databases suffer.

ASFA maintains a uniqueness which acts as a buffer to such competition, but even this database will suffer from the above competition. Google Scholar adds another negative impact because although it may not be as good as ASFA (or the above publishers' products), it may be 'good enough'. Combined with a favorable price – \$0 – traditional users of ASFA may find Google Scholar attractive.

It is imperative that the ASFA Partnership continues to input unique information, whether as part of their normal responsibilities, or as a result of an ASFA Trust Fund proposal. The input of 'regular' input – records from journal articles – must be accurate and timely. Although timeliness has improved overall, there are still some significant publications (core journals) which are only available to ASFA users after several months. Timeliness must improve, and the focus of our efforts must move towards unique content, as this will ultimately determine the long-term viability of the ASFA database.

CSA Action Items Arising from the 2004 ASFA Board Meeting

“Action CSA 2. Regarding the recruitment of an ASFA Partner in the Netherlands

It was suggested that knowing where Dutch institutes carried out their bibliographic searches and which Dutch institutes were subscribing to ASFA could be useful in identifying other potential ASFA Partners in the Netherlands.

CSA agreed to try to provide such information. (section 6.2.2, pg. 4)”

Resolution: Done (e-mail to Richard Pepe, July 2004)

3. Renewed (2004-2007) ASFA Publishing Agreement between FAO and the ASFA Publisher (CSA)

Referring to Exhibit F, Item 5 of the Publishing Agreement (Annex-42) some Partners questioned the limited number (four) of additional ASFA entitlements that Partners may claim after a certain number of input sheets are submitted to CSA.

Agreed CSA agreed to consider increasing the number of IDS subscriptions granted to Partners on a case-by-Action CSA case basis. (section 6.4, pg. 6)

Resolution: Done (July 2004)

7. Regarding Automatic Data Entry

Action CSA CSA agreed to send FAO details of the software (spider and Visual Basic programs) used to download and process citations from one Publisher (and how/where to get it and provide a text file with the Visual Basic program). (section 7.5.1, pg. 12)

Resolution: Attached (Annex 5); initiated several discussions with partners as to the details

9. Machine-assisted indexing

Action CSA CSA agreed to provide FAO with a selected list of records from all ASFA Partners (e.g. 100) which would be posted via FTP, so that one can see the same records indexed both manually and by machine. Should ASFA Partners also wish to receive individual lists of records, they should E-mail their request to CSA. (section 7.5.3, pg. 13)

Resolution: MAI'd records presented to individual centres at their request (e.g. United Kingdom). Note: there is little direct advantage in the comparison as no records are released without an editor reviewing every abstract record, and modifying as appropriate. A comparison may be of general interest to get an idea of how the software works.

11. Regarding the non-renewal of the contract between CSA and SilverPlatter/Ovid regarding the production of the ASFA CD-ROM

Action All 11.1 The ASFA Board agreed that they would test the new ASFA CD-ROM produced by CSA and send their comments to FAO, who would then notify CSA. (section 8.2, pg. 14)

Action CSA 11.2 CSA agreed to provide ASFA Partners with instructions on the setting up of the new CD-ROM on LAN, following a request by Mr Sainekar (NICMAS). (Rapporteurs note: this had been done) (section 8.2, pg. 14)

Resolution: Done

Action CSA 11.3 CSA agreed to investigate the possibility of making the CD-ROM Linux-compatible. (section 8.2, pg. 14)

Resolution: Done (see comment above regarding status of CD).

12. Regarding the ASFA on the CSA Internet Service

Action CSA CSA agreed to provide the Partners with a list of further improvements/changes that will be made to the IDS in the future, and mentioned that some improvements are already listed in the CSA Report. (section 8.3, pg. 15)

Resolution: Done (see comments in the 2004 and current report)

15. Regarding the setting up a Web Site as a depository for electronic documents, in particular for grey literature documents

Action FAO 15.1 The FAO ASFA Secretariat agreed to further consider/study the matter including the possibility of using its Web Server to set up a Web Site for electronic documents, where all the grey literature contained in the various individual ASFA Partners repositories could be kept together and linked to the ASFA database. (section 8.4, pg. 17)

Action CSA 15.2 CSA agreed to provide FAO with the technical details of what would be involved in setting up a Web server. (section 8.4, pg. 17)

Resolution: In process. We are currently experimenting with the conversion and display of 1,000 page Environmental Impact Statement reports to determine the exact costs of hardcopy digitization, and cost of display and maintenance.

18. Regarding Partners still using the DOS- ASFISIS (release-3) software

Action CSA CSA agreed to provide FAO information as to who was still preparing ASFA input using the ASFISIS and FAO software, and the FAO ASFA Secretariat agreed to contact those ASFA Partners to investigate "why" (if there was a technical problem) and try to get them to change to the new www-ISIS-ASFA software. (section 9.1 pg. 18)

Resolution: Done

20.1 ASFIS-1, Serials Monitored for the ASFIS Bibliographic Database

Action CSA CSA agreed to send FAO a list of publications with some information regarding publication dates (i.e. when records from a particular journal first start appearing in ASFA). (section 11.1 pg. 19)

Resolution: Not done. These data are now available and will be sent to FAO as soon as possible.

Annex 1: ASFA Partners' Input (1987-Sept. 2005)

Centre	Code	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993	1992	1991	1990	1989	1988	1987
ADRIAMED	AA	60	34	136	171	133	60													
Argentina	AR	60	53	62	53	39	85	68	60	74										
Australia	AU	135	63	221	125	349	297	261	70	0	0	896								
Belgium	BE		280	97	77															
Brazil	BR		12																	
Bulgaria	BG																			
Canada	CA	213	595	590	806	881	737	951	1,052	1,095	846	1,233	1,096	1,034	919	927	947	1,010	986	889
Chile	CL		20	49	0	0	0	0	39											
China	CH	667	3,700	2,312	2,320	2,409	1,198	1,163	1,205	1,181	1,078	1,191	1,060	1,042	705	1,022	1,007	1,022	1,010	1,004
Cote d'Ivoire	CI		20	0	0	63	20													
Cuba	HV		44	78	21	0	16	33	17	0										
Estonia	EE	13	115	61	22	42	38	48	46	63	33									
FAO	FA	1,122	742	1,053	1,731	1,188	1,851	1,968	1,628	1,414	1,203	2,195	2,895	2,205	3,801	2,682	2,841	4,020	3,870	4,379
France	IF	1,068	1,280	1,253	1,232	1,175	1,651	1,769	1,591	1,406	1,375	1,296	1,121	1,111	1,308	1,100	950	1,111	1,390	1,599
Germany	BF	332	787	1,164	1,187	1,876	1,917	2,105	2,294	1,990	2,008	2,000	1,738	2,000	2,000	1,904	1,956	1,700	1,702	2,033
Greece	GR	101	332	253	250	230	160	56	24	20										
ICCAT	TU	505	0	0	0	91	16													
ICES	CE	28	283	593	1,013	96	65	86	61	46										
ICLARM	IC		55	67	251	0	0	42	125	239	514	10								
IMROP	MR	19																		
India	DP	501	783	991	1,073	1,050	798	1,541	967	807	870	770	866	841	795	854	817			
IOC	IO	28	22	113	77	109	97	44	96	637	19									
IOC-HA	HA				0	0	0	312												
Iran	IN	14																		
Italy	IT	95	53	402	102	168														
IUCN	IU		34	44	58	0	0	174	39	55	41	28	25							
Japan	JA		974	1,391	971	362	808	805	712	639	596	644	706	691	662	656	640	668	687	719
Kenya	KE	102	167	124	133	715	606	70	126	38	30									
KORDI	KR				250	450	50													
Lithuania	LT				0	0	0	119	71	127	15	122	101	112	43					
Mexico	MX		312	357	552	495	629	334	699	34	440	702	720	604	522	770	490	814	771	307
Morocco	MO																			
Nigeria	NG	107	143																	
Norway	SK	166	391	437	255	269	570	286	303	476	254	95	0	0	154	134	50			117

Peru	PE		104																	
PIMRIS	PI			88	11	30														
Poland	PL	159	210	188	116	87	113	116	61	62	48									
Portugal	PT	36	38	0	0	0	157	55	0	107	120	170	63	49	61	70	199	351	219	167
Russia	VN	183	170	162	307	371	300	328	239	877	581	476	628	852	676	368	700	658	880	975
Sec. Pacific Comm.	CP	56	5																	
SIPAM (FAO)	SP			16	16	18	30													
Spain	ES				0	138	106	96												
Sweden	SE	41	90	61	85	305	27	79	44	49	37									
Tunisia	TN	34	42	0	28															
UK	MB	1,247	1,609	1,212	1,288	1,473	1,215	1,639	1,264	1,030	785	1,079	994	1,300	610	1,400	1,200	925	1,000	1,000
Ukraine	UA	36	201	236	290	289	100	180	140	150	316	634								
UN	UN	102	204	231	172	195	195	184	199	197	160	138	113	117	161	139	153	100	160	81
UNEP	EP	101	0	100	0	0	0	0	0	0	0	0	82	136	80					
Uruguay	UY	40	60	40																
USA	NO	1096	1455	1,918	1,803	2,359	1,931	2,399	2,810	3,195	3,784	3,913	3,019	3,258	1,800	2,213	2,430	1,943	1,920	1,838
Viet Nam	VI			81																
CSA	CS	16481	22773	23,793	24,875	24,799	21,873	25,211	25,826	23,486	23,705	26,274	22,866	21,168	20,531	19,870	17,703	14,726	10,319	12,697
Total		24,948	38,255	39,974	41,721	42,254	37,716	42,522	41,808	39,494	38,858	43,866	38,093	36,520	34,828	34,109	32,083	29,048	24,914	27,805

Annex 2: ASFA Paid Subscriptions by Country (2004)

(CSA Illumina, Print; & CD Sales. Note that CSA Illumina sales are for ASFA-only products, and does not reflect number of database clusters sales which include ASFA as a component.)

Country	CSA Illumina					Print					NISC	Total
	ASFA-1	ASFA-2	ASFA-3	ASFA-Aq	ASFA-Set	ASFA-1	ASFA-2	ASFA-3	ASFA-Aq	Print Set		
Algeria										1		1
Australia	4	3	18							1	4	30
Austria								1				1
Bahrain					1							1
Bangladesh						1						1
Belgium						1						1
Brazil					4					1	1	6
Brunei								1				1
Canada	2	1	1		15					4	3	26
Chile				2	3							5
China					3	1				1	1	6
Colombia					1			1				2
Costa Rica										1		1
Croatia						1	1				1	3
Czech Rep.					1							1
Finland					3							3
France	1				7	4		1		7	1	21
Germany					9	1		1		1	9	21
Greece						1	1				1	3
Greenland					1							1
Hong Kong					2						1	3
Hungary	1		1			1					1	4
India					2	7		2		3	6	20
Ireland					1	1						2
Israel					5							5
Italy					6	6		3			1	16
Japan	1				2	9		1		3	1	17
Korea											2	2
Kuwait					1					1		2
Malaysia										1	4	5
Malta					1							1
Mexico					6	1						7
Namibia					1						2	3
Netherlands		1			3	1					1	6
New Zealand					8							8
Norway				1	9							10
Oman					1							1
Pakistan					1							1
Philippines											2	2
Poland						1	1					2
Portugal				1	1							2
Puerto Rico											1	1
Qatar					1							1

Russia									1		1	
Singapore										2	2	
South Africa									1	12	13	
Spain	1			17		2	1		1	2	24	
Sweden			1	5							6	
Switzerland				1							1	
Taiwan												
ROC				6							6	
Thailand				1						1	2	
Trinidad & Tobago				1							1	
Turkey				1							1	
UK	1			6					2	10	19	
USA	12	2	1	4	104	17			4	12	156	
Grand Total	23	7	22	8	241	54	5	12	0	34	82	488

Annex 3. ASFA Prices

CSA 2005 Pricing of ASFA Products (US\$)

Title	CSA Illumina	Print	Web Edition	CD
Aquatic Sciences & Fisheries Abstracts (ASFA) Database Set	\$6,200-7,675	\$2,795	2,095	Negotiable
ASFA 1: Biological Sciences & Living Resources	\$3,775	\$1,755	1,315	n/a
ASFA 2: Ocean Technology, Policy & Non-Living Resources	\$2,925	\$1,280	960	n/a
ASFA 3: Aquatic Pollution & Environmental Quality	\$1,025	\$550	410	n/a
ASFA Aquaculture Abstracts	\$1,350	n/a	440	n/a
ASFA Marine Biotechnology Abstracts	\$1,000	n/a	390	n/a

Annex 4. ASFA Marketing Efforts 2005

Dates	Show	City	State	Country
15-17 Jan	ALA Midwinter	Boston	MA	USA
10-13 Jan	American Meteorological Society	San Diego	CA	USA
1-3 Feb	Information Online 2005	Sydney		Australia
3-5 Feb	Ontario Library Association	Toronto	ON	Canada
14-16 Feb	TMS Annual	San Francisco	CA	USA
26 Feb - 1 Mar	Council on Social Work Education	New York	NY	USA
16-18 Mar	Computers in Libraries	Washington	DC	USA
3-5 Apr	Art Libraries Society of North America	Houston	TX	USA
7-8 Apr	Tennessee Library Association	Nashville	TN	USA
7-9 Apr	ACRL	Minneapolis	MN	USA
Apr	SAIL (Southeast Affiliate of IAMSLIC Libraries)	Baltimore	MD	USA
4-5 May	Amigos Member Conference	Dallas	TX	USA
	CALIS (China Academic Library & Info Sys) Consortium			China
May	AMBAC (Asociación Mexicana de Bibliotecarios, A.C.)			Mexico
	iGroup Annual Users Meeting			
27-28 May	Atlantic Provinces Library Association	Halifax	NS	Canada
May	ACURIL			
5-7 Jun	SLA	Toronto	ON	Canada
7 Jun	MOBIUS Vendor Fair	Lake of the Ozarks	MO	USA
13-15 Jun	ASEE	Portland	OR	USA
25-28 Jun	ALA	Chicago	IL	USA
Jun	Scientific Technological Information Management Assoc (STIMA)			Korea
Jun	COMPAB			Mexico
13-16 Aug	American Sociological Association	Philadelphia	PA	USA
14-17 Aug	IFLA	Oslo		Norway
1-4 Sep	American Political Science Association	Washington	DC	USA
Sep	LIANZA			New Zealand
Sep	Korea Library Association			Korea
Oct	IAMSLIC	Rome		Italy
	CONACYT (Consejo Nacional de Ciencia y Tecnología)			Mexico
	Hawaii Library Association			USA
	LOUIS Consortium		LA	USA
	JULAC (Joint University Librarians Advisory Committee)	Hong Kong		China
	NYLINK		NY	USA
24-26 Oct	Internet Librarian	Monterey	CA	USA
	CONCERT 2005 (Consortium on Core Electronic Resources in Taiwan)			Taiwan
Nov	Charleston Showcase	Charleston	SC	USA
Nov	Library Fair			Japan
29 Nov-1 Dec	Online Information 2005	London		UK
Dec	Military Librarians Workshop			

Annex 5. Procedure for downloading Web documents

Semi-automatic Downloading of Web Documents:

CSA obtains abstract records from several thousand journals by downloading and tagging html files from publishers web sites (e.g. Science Direct, Wiley, Springer, Kluwer, Blackwell, NRC-Canada, etc.). Although some are formally under the control of our IT department, requiring specific programmers, there is a relatively easy way to download and process records that do not require extensive programming knowledge. The following describes the general nature of the downloading process:

Two types of spiders/crawlers/screen scrapers/robots are used to download HTML-coded files from publishers' web sites.

1. Macro Software (e.g. those that reproduce key strokes such as www.macroexpress.com)
 - This macro uses a citation document, listing all the volumes and issues expected from particular journals, and reproduces key strokes that take that information to identify web pages containing the abstract records for particular journals.
2. Crawlers (e.g. the wholesale downloading of every link from a publisher web site (e.g. www.webreaper.net).
 - This software is indiscriminant; you have to select the 'real' files from the chaff at a publisher web site.

Either of the above two methods provides you with single html files of individual abstracts. These files often take the following form:

```
<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
<!-- TRANSKEY: 09/27/2005 10:30:24 EDT#1861_004#33409#199.1.202.75 -->
<title>ScienceDirect - Acta Biomaterialia :
Three-dimensional co-culture of rat hepatocyte spheroids and NIH/3T3 fibroblasts enhances hepatocyte functional maintenance</title>
</head>
<link rel="stylesheet" Title="Default Styles" href="http://www.sciencedirect.com/science/page/static/stylesheets/css_txt_v01.css" type="text/css">
<body bgcolor="#ffffff" text="#000000" >
<!--
...
1438e8fb51" target="_top">Privacy Policy</a>
</font>
</font></td></tr>
<tr><td height="10"></td></tr>
</tr></table>
</body>
</html>
```

After the individual files are combined (e.g. www.txtcollector.com) into a single file, the html files are run through MS Word macros which replace key html tags with the appropriate tags to load into our production system (or any production system, such as Web ISIS). At CSA, these files typically look like the following:

```
ISSN: 0960-8524
JI: 0960-8524, |v 96 |n 14 |p |s 20050900 |f |r 20050609
PI: 1551-1560
DOI: 10.1016/j.biortech.2004.12.016
TI: A diffusion model for drying of a heat sensitive solid under multiple heat
input modes
AU: Sun, L.
AU: Islam, M.
AE: mpeasm@nus.edu.sg
AF: Department of Mechanical Engineering, National University of Singapore, 10
Kent Ridge Crescent, Singapore 119260, Singapore
AB: To obtain optimal drying kinetics as well as quality of the dried product in
a batch dryer, the energy required may be supplied by combining different modes
of heat transfer. In this work, using potato slice as a model heat sensitive ...
```

Any special characters or formatting codes are translated to standard CSA codes.