

## **Open Access – Open Source - szabad hozzáférés tudományos cikkekhez, szabad szoftverrel**

NIIF IPSZILON szeminárium, 2007 június 19.

Csirmaz L., Holl A.

### ***Összefoglalás***

A tudományos eredmények szabad hozzáférhetősége – ez lényegében a szakfolyóiratokban megjelent cikkek hozzáférhetőségét jelenti – nagy jelentőséggel bír. Az Internet lehetőséget ad erre: a kutatók közölhetik cikkeiket elektronikus formában szabadon elérhető folyóiratokban, és lehetőségük van a hagyományos folyóiratokban publikált cikkek kéziratának szabadon hozzáférhető lerakatokban: repozitóriumokban való elhelyezésére. Az OpenAccess mozgalom kidolgozott egy protokollt a szabadon elérhető szakcikkek láthatóvá tételére: ez az Open Archives Initiative – Protocol for Metadata Harvesting. Mind az elektronikus folyóiratok, mind a repozitóriumok üzemeltetésére többféle szabad szoftver áll rendelkezésre.

---

### Vázlat

- Az Open Access mozgalom: kezdeményezések a tudományos szakirodalom szabadon hozzáférhetővé tételére; az Open Access jogi lehetőségei (Csirmaz L.)
- Elektronikusan hozzáférhető szakirodalom ma; bibliográfiai szolgáltatások; szabad elérésű szolgáltatások a fizika és a csillagászat területén; az Open Archives kezdeményezés és az OAI-PMH; (Holl A.)
- Egy szabad forráskódú repozitórium szoftver: az EPrints (Csirmaz L.)
- Egy szabadon hozzáférhető elektronikus újság: az IBVS; az IBVS-t működtető szoftverek bemutatása; a YAR statikus repozitórium. Recept elektronikus újságok ingyenes szoftverrel való üzemeltetésére. (Holl A.)

### ***Bevezetés és igehirdetés***

(Lásd Csirmaz L. vetített anyagát)

- Tudomány: publikálni kell, az eredményesség mértéke az idézettség.
- Jelenlegi helyzet: tudományos folyóiratok, üzleti alapon működő kiadóvállalatok, magas (és egyre növekvő) előfizetési díjak, amelyek az elektronikus forma megjelenése után sem csökkentek, sőt... A kiadók állítása: a magas minőség biztosítása, az innováció költségei [és

a reklám költségek] miatt nem lehet mindez olcsóbb. Ismerős? Ellenvélemény: az adófizetők pénzén létrehozott tudományos eredményeket miért kell az adófizetők pénzén visszavásárolni?

- A megoldás: BOAI (és a többi kezdeményezés); két út: arany (a weben ingyenesen elérhető folyóiratok: Holl A.) és zöld (hagyományos folyóiratokban közölt cikkek elhelyezése repozitóriumokban: Csirmaz L.).

Jó a közösségnek, de miért erőltesse meg magát a szerző vagy a munkáltatója, kutatásának finanszírozója? Egyrészt: ha ők nem, más miért tegye? Másrészt: a megnövekedett láthatóság, több idézet miatt.

- Vajon az ingyenes folyóiratok, repozitóriumok tényleg nagyobb láthatóságot jelentenek? Hasonló a helyzet az üzleti weblapok esetéhez: minél előbb jelenik meg a keresőkben, minél több link mutat rá, annál nagyobb az üzlet (annál több az idézet). Tudományos publikációknál ez a tudományos bibliográfiai rendszerekben való megjelenést és a referencialinkeket jelenti. Biztosítja ezt a szabad hozzáférés (OA)? Igen! (statisztika).

---

---

### ***Helyzetkép: szolgáltatások, technológiák***

Holl A.

- Komersz valóság: a tudományos folyóiratok elektronikus változatai

Elektronikus formában általában a 90-es évek végétől fogva léteznek az egyes folyóiratok. Visszamenőleg ritkán érhetőek el az egyes számok elektronikusan, ha igen jobbra csak PDF verzióban. Az elektronikus változatok a PDF formátumú verziók mellett gyakorta HTML változatot is kínálnak, esetleg strukturálva (cikk tartalomjegyzék), keresési lehetőségeket nyújtanak (keresés a meta-adatokban). Gyakran alkalmazzák a referencia-linkkeket. Az egyes cikkek nem minden esetben címezhetőek, érhetőek el közvetlenül.

Vannak olyan elektronikus szolgáltatások, melyek több kiadó folyóirataihoz kínálnak hozzáférést (példa: EBSCO), illetve egy előfizetői konzorcium tagjai számára tesznek hozzáférhetővé folyóiratokat (példa: EISZ).

A bibliográfiai szolgáltatók között megtalálhatóak a régi szolgáltatások elektronikusan elérhető változatai, és vannak új szereplők a piacon. A közvetlen bibliográfiai adatszolgáltatókon keresőrendszeren (Web of Science, Scopus) túl léteznek derivatívák is, mint az idézetek ill. az újságok impakt faktorai (SCI). Az információ begyűjtése többnyire hagyományos módszerekkel történik.

Fontos technológia a minden elektronikusan hozzáférhető dokumentum egyedi azonosítását lehetővé tevő DOI, ill. az erre épülő, referencia-linkelést lehetővé tevő szervezet: a CrossRef. Ennek a szolgáltatásnak az igénybevétele a kiadóknak pénzbe kerül.

- A fizikai tudományok területe (és ezen belül a csillagászat) különleges helyzete

Az átlagosnál többet nyújtó, ráadásul sokszor ingyenes szolgáltatások léteznek. A 90-es évek elejétől létezik a szabadon hozzáférhető fizikai kézirat (preprint) lerakat (jelenleg: arXiv).

A csillagászatban kialakult a DOI-t megelőzően egy egységes bibliográfiai azonosító: a bibcode (<http://cdsweb.u-strasbg.fr/simbad/refcode/refcode-paper.html>).

A NASA és a Harvard-Smithsonian Center for Astrophysics fenntart egy bibliográfiai rendszert, az ADS-t. Nem csupán bibliográfiai szolgáltató, de visszamenőlegesen digitalizálta a csillagászati szakirodalom jelentős részét. Fejlett funkcionalitást kínál pl. hasonló tartalmú cikkek keresésére, ill. szerzői nevek azonosítására. Létezik két nagy csillagászati adatbázis (CDS SIMBAD és a NED), melyekben az objektumokhoz bibliográfia is tartozik.

Léteznek ingyenesen elérhető folyóiratok (pl. IBVS), csak elektronikusak (PASA), a legfontosabb folyóiratok az átlagot meghaladó szolgáltatásokat nyújtanak, valamint elterjedt a cikkekhez kapcsolódó megfigyelési anyagok elérhetősége.

– Szabad szoftverek, szabad hozzáférés

Az Open Archives Initiative létrehozta a Protocol for Metadata Harvesting-et (OAI-PMH). A protokoll HTTP alapú, és DublinCore szabványú meta-adatokat terjeszt XML formátumban. Ez biztosítja szabad, szabványos, korszerű módon a folyóiratcikkek láthatóságát, a meta-adatok elérhetőségét. A teljes értékű repozitóriumok (lerakatok) mellett, kisebb méretű és forgalmú gyűjteményekhez statikus repozitóriumokat is létre lehet hozni, amelyek egy XML állományból és egy szerverből (Static Repository Gateway) állnak. A PMH adatszolgáltatókat egy központi nyilvántartóba (registry) lehet bejegyeztetni. Volt egy próbálkozás a referencia-linkelés megoldására: az OpCit. Az arXiv után már számos szabadon hozzáférhető kézirat lerakat létezik – de ezek még nem képviselnek átütő erőt. Az OAI-PMH alapokra rengeteg szabad keresőszolgáltatás, bibliográfiai rendszer épül (CiteSeer, citebase, OAIster, DOAJ, stb.), de ezek még nem kiforrottak. (Bár a kommersz bibliográfiai keresők használhatósága sem mindig jó. Kivételes az ADS a csillagászatban, mely ráadásul ingyenes!) Kereskedelmi vállalkozásban, de ingyenesen elérhetően működik a Google Scholar. Ez nem a meta-adatok szabványos protokollal való cseréjén, hanem a Google webjáró robotjain alapul. A jövő véleményünk szerint az OAI-PMH. De számos kérdést meg kell még oldani: pl. bővebb meta-információk kellene a DublinCore által nyújtottnál és meg kell oldani a redundáns tételek kiszűrését.

– Szabad szoftverek

Léteznek szabad repozitórium szoftverek, és olyan szabad szoftverek, amelyek az OAI-PMH protokollú meta-adat terjesztésén túl az elektronikus újságok munkafolyamat támogatását is elvégzik, és webes felületet nyújtanak a folyóirat számára.

- Szabványos formátumok

Figyelmet kell fordítani a szabványos és nyílt formátumokra. A legfontosabb szempont az időállóság, ez pedig szabványosságot kíván. Nem hanyagolható el az az igény sem, hogy a szakirodalmat szabad szoftverekkel lehessen olvasni. (Példák: SGML, XML, HTML, ODF, PDF/PostScript, TeX/LaTeX.)

---

---

### ***Repozitóriumok***

(Lásd Csirmaz L. vetített anyagát)

- OSI guide
- 
- 

### ***Független, szabad elérésű tudományos folyóiratok***

Holl A.

- Üzleti modellek: Valaki mindig fizet – legfeljebb kevesebbet. OSI publikálási üzleti segédlet. Egy lehetőség: a PLoS (szabadon elérhető elektronikus folyóirat, OSI támogatás a publikálásra). Szabad elérésű folyóiratok / szabad forráskódú szoftverrel üzemeltetett folyóiratok – költségcsökkentés lehetséges.
- SPARC információs oldal: kommersz és szabad folyóiratüzemeltető szoftverek, stb. (A folyóiratüzemeltetésnek és publikálásnak számtalan olyan oldala van, ami más területeken alkalmazott szoftverek (pl. web portál építő, csoportmunka) használatával is megoldható.)  
<http://www.arl.org/sparc/resources/pubres.html#journals>
  - Open Journal System  
Public Knowledge Project, Univ. of British Columbia / Simon Fraser Univ.  
nyílt forráskódú (GPL) újságüzemeltető és publikációs rendszer  
UNIX, MySQL/PostgreSQL, PHP  
<http://pkp.sfu.ca/?q=ojs>
  - Zope  
<http://www.zope.org/>  
Nyílt forráskódú alkalmazáserver tartalomkezelő rendszerek, portálok  
és egyéb alkalmazások építésére  
Python

- LIVING REVIEWS ePublishing Toolkit

Max-Planck-Gesellschaft

<https://dev.livingreviews.org/projects/epubtk/>

Unix, Windows; Python, PostgreSQL, Apache, TeX,..

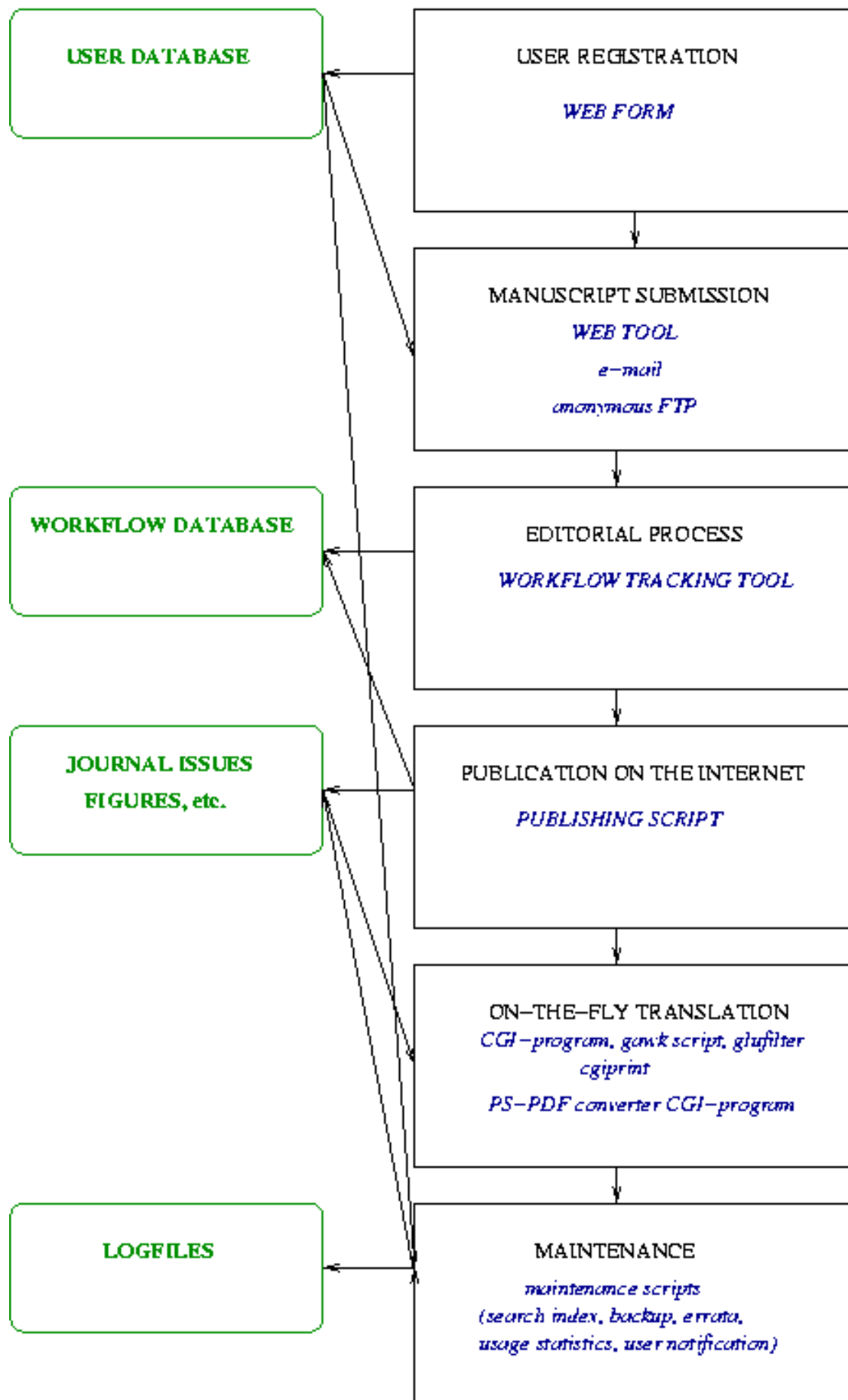
példa újság: Living Reviews in Solar Physics

<http://solarphysics.livingreviews.org/>

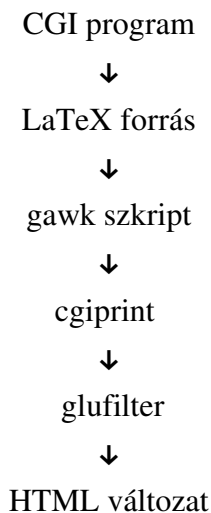
- Egy egyszerű webes folyóiratváltozat: a Magyar Tudomány  
<http://www.matud.iif.hu/>
- Technikai feltételek az adatbázisokban való indexelhetőségre, keresztlinkelésre
  - egyedi azonosító (DOI, bibcode vagy hasonló)
  - cikkek közvetlen címezése (URL, lehetőleg az egyedi azonosító alkalmazásával)
  - meta-adatok
- Csináld magad: az IBVS (és a csillagászat speciális helyzete)  
Az Information Bulletin on Variable Stars egy erősen specializált szakfolyóirat, amelyet az MTA Konkoly-Thege Miklós Csillagászati Kutatóintézete ad ki a Nemzetközi Csillagászati Unió (IAU) 27-es és 42-es Kommisszióinak megbízásából. 1994/95 óta szabadon elérhető az Interneten.

Az IBVS készítésének munkafolyamata:

# IBVS workflow chart



- HTML cikkek szolgáltatásai
- markup (LaTeX)
- HTML készítés reptében (ideális esetben minden elkészíthető a LaTeX forrásszövegből és az ábrákból, adatállományokból)



- adatbázis-linkek (SIMBAD, GCVS)
- URL-ek, e-mail címek (elrejtés adatvédelmi okból), Vizier és arXiv linkek
- referencia-linkek: szövegértelmezés a gawk szkript segítségével
- fejlett technológia: a CDS Aladin használata grafikus eszközként – GRID technológia
- Az IBVS láthatósága
  - meta-információk – ADS tagged format, DublinCore XML
  - hagyományos módszer (push): e-mail az ADS-nek és a CDS-nek
  - új, OAI-PMH alapú (pull): YAR (<http://yar.sourceforge.net/>)  
A YAR statikus repozitórium/és webes PMH szolgáltató.  
Perl alapú, egyrészt olyan programokat tartalmaz (tapecreate, tapeindex), amelyekkel egy statikus XML állományba (tape) lehet a meta-adatokat összeállítani, másrészt egy Apache alá illeszthető PMH szolgáltatást (oai2.pl). Kiválóan illeszkedik az IBVS programrendszerébe: OAI-PMH támogatást ad, és semmi mást.
  - kitekintés: miért szükségesek a saját meta-információk? Google Scholar robotjai
- Az IBVS üzemeltetése: webes cikkbeküldő felület, szerkesztőségi információs rendszer, lapműködtetés (folyamatábra)

#### *Recept saját kiadású elektronikus folyóirat üzemeltetéséhez*

- Nem technikai szükségletek
  - Min. két státusznak megfelelő munkaerő
  - (Nemzetközi) kapcsolathálózat bírálók igénybevételére
  - Lapgazda szervezet megfelelő presztízzsel, stabilitással

- Technikai szükségletek
    - Szabványos formátum
    - Markup (meta-adatok kiemelése, lehetőleg a szerzők által)
    - Szabad szoftver a lap üzemeltetésére
    - OAI-PMH protokoll megvalósítása a folyóirat láthatóvá tételére
- 

## YAR telepítés

- Perl csomag, tömörített tarfile

CPAN-nal? A telepítési útmutatóban benne van a megkívánt modulok listája. Fordítás/installálás: `make`.

- Apache konfigurálás:

```
ScriptAlias /yar /usr/local/bin/oai2
```

- DublinCore XML meta-adat állományok készítése:

`dcmetas.csh` – c-shell wrapper ; `IBVSdc.awk` – konvertáló program

input: ADS formátumú meta-adat állományok

- „tape” (XML állomány) készítés: `tapecreate`, `tapeindex`

- módosítani kell az `/etc/yar/oai2.conf` állományt

## Használat:

```
http://buda.konkoly.hu/yar/ibvs?verb=Identify
http://buda.konkoly.hu/yar/ibvs?verb=ListMetadataFormats
http://buda.konkoly.hu/yar/ibvs?verb=ListIdentifiers&metadataPrefix=oai_dc
http://buda.konkoly.hu/yar/ibvs?verb=ListRecords&metadataPrefix=oai_dc
```

---

## *URL-ek, irodalom*

IBVS - Information Bulletin on Variable Stars

International Astronomical Union Commissions 27 and 42

<http://www.konkoly.hu/IBVS/IBVS.html>

MTA Konkoly-Thege Miklós Csillagászati Kutatóintézete

Konkoly Observatory of the Hungarian Academy of Sciences

<http://www.konkoly.hu>



SCI - Thomson Science Citation Index

Thomson Scientific

<http://www.thomsonscientific.com/cgi-bin/jrnlst/jloptions.cgi?PC=K>

Web of Science

<http://scientific.thomson.com/products/wos/>

Science Direct

Elsevier

<http://www.sciencedirect.com>

Scopus

<http://www.scopus.com/scopus/home.url>

EBSCO Academic Search Premier

<http://search.ebscohost.com/>

EISZ – Elektronikus Információszolgáltatás

<http://eisz.om.hu>

DOI

<http://www.doi.org/>

CrossRef

<http://www.crossref.org/>

BOAI - Budapest Open Access Initiative

<http://www.soros.org/openaccess/>

OAI - Open Archives Initiative

<http://www.openarchives.org/>

OAI-PMH – Protocol for Metadata Harvesting

<http://www.openarchives.org/pmh/>

<http://www.openarchives.org/OAI/openarchivesprotocol.html>

<http://hektar.sztaki.hu/oai/protokoll.html>

Static Repository & Static Repository Gateway

<http://www.openarchives.org/OAI/2.0/guidelines-static-repository.htm>

OAI registry

<http://www.openarchives.org/data/registerasprovider.html>

OpCit - Open Citation Project

<http://opcit.eprints.org/>

citebase

Tim Brody, Univ. Southampton

<http://www.citebase.org/>

Google Scholar

<http://scholar.google.com/>

CiteSeer

Penn State Univ.

<http://citeseer.ist.psu.edu/>

DOAJ – Directory of Open Access Journals

<http://www.doaj.org/>

Lund University Libraries

SHERPA - Securing a Hybrid Environment for Research Preservation and Access

<http://www.sherpa.ac.uk/index.html>

RoMEO

<http://www.sherpa.ac.uk/romeo.php>

OpenDOAR – Directory of Open Access Repositories

<http://www.opendoar.org>

OAIster

<http://www.oaister.org/>

University of Michigan

ScientificCommons

„access to free scientific knowledge”

<http://en.scientificcommons.org/>

Univ. of St. Gallen

BASE - Bielefeld Academic Search Engine

[http://base.ub.uni-bielefeld.de/index\\_english.html](http://base.ub.uni-bielefeld.de/index_english.html)

SGML - Standard Generalized Markup Language

Charles Goldfarb, Edward Mosher, Raymond Lorie

<http://en.wikipedia.org/wiki/SGML>

XML – eXtensible Markup Language

Tim Bray

W3C

<http://hu.wikipedia.org/wiki/XML>

HTML – HyperText Markup Language

Tim Berners-Lee

W3C

<http://hu.wikipedia.org/wiki/HTML>

ODF - OASIS Open Document Format for Office Applications

<http://en.wikipedia.org/wiki/OpenDocument>

PDF - Portable Document Format

Adobe

[http://hu.wikipedia.org/wiki/Portable\\_Document\\_Format](http://hu.wikipedia.org/wiki/Portable_Document_Format)

PostScript

<http://en.wikipedia.org/wiki/PostScript>

TeX

Donald Knuth

<http://en.wikipedia.org/wiki/TeX>

LaTeX

Leslie Lamport

<http://www.latex-project.org/>

<http://en.wikipedia.org/wiki/LaTeX>

OSI Guide to Open Access Publishing and Scholarly Societies

[http://www.soros.org/openaccess/scholarly\\_guide.shtml](http://www.soros.org/openaccess/scholarly_guide.shtml)

<http://www.soros.org/openaccess/oajguides/>

OSI Guide to Institutional Repository Software

<http://www.soros.org/openaccess/software/>

PLoS - Public Library of Science

<http://www.plos.org/>

Independent Journal Program, BioMed Central

BioMed Central Ltd.

<http://www.biomedcentral.com/home/>

DublinCore

<http://dublincore.org/documents/dc-xml-guidelines/>

<http://dublincore.org/documents/dces/>

SPARC - The Scholarly Publishing & Academic Resources Coalition

<http://www.arl.org/sparc/>

SPARC lista

<http://www.arl.org/sparc/resources/pubres.html#journals>

YAR - Yet Another (OAI-PMH) Repository

LANL

<http://yar.sourceforge.net/>

arXiv.org

P. Ginsparg

LANL/Cornell Univ. Library

<http://arxiv.org/>

ADS - Astrophysics Data System

NASA/Harvard-Smithsonian Center for Astrophysics

<http://www.adsabs.harvard.edu/>

ADS tagged meta-information format

[http://doc.adsabs.harvard.edu/abs\\_doc/help\\_pages/taggedformat.html#taggedformat](http://doc.adsabs.harvard.edu/abs_doc/help_pages/taggedformat.html#taggedformat)

SIMBAD

Centre de Données astronomiques de Strasbourg

<http://cdsweb.u-strasbg.fr/CDS.html>

<http://simbad.u-strasbg.fr/simbad/sim-fid>

Aladin Sky Atlas

<http://aladin.u-strasbg.fr/>

VizieR

<http://vizier.u-strasbg.fr/viz-bin/VizieR>

glufilter

P. Fernique

<http://simbad3.u-strasbg.fr/glu/glu.htx>

NED

NASA/IPAC Extragalactic Database

<http://nedwww.ipac.caltech.edu/>

bibcode

<http://cdsweb.u-strasbg.fr/simbad/refcode/refcode-paper.html>

PASA - Electronic Publications of the Astronomical Society of Australia

<http://www.atnf.csiro.au/pasa/>

<http://www.publish.csiro.au/?nid=138>

Holl András: Elektronikus folyóiratok - lehetőségek, technológia, problémák

TMT 2000/9-10

[http://tmt.omikk.bme.hu/show\\_news.html?id=1480&issue\\_id=27](http://tmt.omikk.bme.hu/show_news.html?id=1480&issue_id=27)

Holl András: Az Information Bulletin on Variable Stars (egy csillagászati szaklap a webkorszakban), 2001, Polaris Csillagvizsgáló

<http://www.konkoly.hu/staff/holl/mcseibvs.html>

Holl András: Elektronikus folyóiratok a természettudományok területén - egy hazai példa

Networkshop, 2001

<http://www.konkoly.hu/staff/holl/sopron/sopron.html>

Hernád István: A tudományos szakirodalom szabaddá tétele az Interneten -

Intézmény-alapú önarchiválási kezdeményezés

Magyar Tudomány, 2002/2

<http://www.matud.iif.hu/02feb/hernad.html>

Bánhegyi Zsolt: Nyílt Hozzáférés Kezdeményezés (Open Access Initiative) - Kitekintés és körkép

TMT, 2003/6-7

[http://tmt.omikk.bme.hu/show\\_news.html?id=2093&issue\\_id=66](http://tmt.omikk.bme.hu/show_news.html?id=2093&issue_id=66)

Yiotis, Kristin: A nyílt hozzáférés kezdeményezés: a tudományos kommunikáció új paradigmája

TMT, 2006/3 (Bánhegyi Zs. ismertetése)

[http://tmt.omikk.bme.hu/show\\_news.html?id=4359&issue\\_id=470](http://tmt.omikk.bme.hu/show_news.html?id=4359&issue_id=470)