The Virtual Library in Action

I

By Louise Addis et al.

Presented at American Chemical Society (ACS) National Meeting Chemical Information Division Symposium (CNF), The Library of the Future, April 1995, Anaheim, CA, USA

Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309

Work supported by Department of Energy contract DE-AC03-76SF00515.

File	Postscript	Navigate	Customize	Documents	Manuals

Document: [http://slacvm.slac.stanford.edu/FIND/slac.html

THE VIRTUAL LIBRARY IN ACTION

The Stanford Linear Accelerator Center (SLAC) provides universal full-text access to the literature of particle physics via the SPIRES-HEP database and the World-Wide-Web (WWW)

L. Addis, H. Galic, P. Kreitz, A. Johnson Stanford Linear Accelerator Center* addis@slac.stanford.edu

Presented at the American Chemical Society (ACS) National Meeting Chemical Information Division Symposium (CINF), The Library of the Future Anaheim, CA

4 April 1995

* Supported by the U.S. Dept. of Energy, Contract DE-AC0-3-76SFO-0515

Keyword:	
Go Back Previous Next · Save Search · Clone Close Window	
	hannan hanna ha

Help

ABSTRACT

The SLAC Library has for many years provided SPIRES-HEP, a 300,000 record bibliographic database, to the world particle physics community via the Internet as well as through clone sites in Europe and Japan. The 1991 introduction of the e-print archives at LANL coupled with the World-Wide-Web (WWW) from CERN suddenly made it possible to provide easy linkage between bibliographic database records and the actual full-text of papers. The SLAC Library has turned this possibility into reality by converting hundreds of TeX source documents each month into viewable postscript complete with figures. These (now more than 20,000) postscript files are linked to the HEP database, and the full-text is rendered universally visible via WWW. We discuss the project, the collaboration of physicists and librarians, what is easy, what is hard, and our vision for the future.

SLAC - HEP Information Services April 1995

Today SLAC supports several databases of interest to the HEP community, including SPIRES-HEP, the HEPnames Email directory (24K addresses), Books, Conferences, Experiments (w/PDG), Institutions, and the FreeHEP software tools directory.

The following diagram shows participation and usage for the SPIRES-HEP database only.





-Sites which run clone copies of the SLAC HEP database and provide regional HEP information. (SLAC provides nightly updates on the Internet. About 20,000 new records are added annually.)
-Sites which use SLAC HEP directly to manage their own document collections.
-Sites which regularly download large portions of SLAC HEP for use in regional services.
-Sites which also contribute data to SLAC HEP or provide other database information.

-Internet user communities.

-Login user communities.

PREPRINTS, E-PRINTS, AND SPIRES-HEP

Preprints (including electronic preprints), conference papers, technical reports, theses, and neverpreprinted journal articles are all entered in the HEP database. Currently about 20,000 new records are added annually by the SLAC Library (60%), the DESY Library (30%), and a variety of other sites. (Like many other aspects of particle physics, SPIRES-HEP is a collaboration.)

The SPIRES-HEP database is tailored to the particle physics community needs thru the inclusion of the usual bibliographic information and some less usual information special to particle physics. For example:

1. All authors are included (with their associated institutions). Since experimental particle physicists tend to work in very large groups on very large and expensive equipment, almost 200 important papers in 1994 had more than 100 authors. The record is 850.

Since authors are almost always listed in alphabetical order, it is extremely important that all authors be searchable (especially to those whose names begin with letters after A, such as Nobelists Rubbia and Richter).

- 2. The names of 'collaborations' are included. Often a group of experimentalists will go by a name based on their experiment such as the SLD Collaboration at SLAC or the CDF and D0 Collaborations at Fermilab which just announced the discovery of the 'top quark.'
- 3. The names of the experiments are included so that all the papers from a particular experiment can be identified.
- 4. All the references to published journal articles or to numbered electronic preprints (e-prints) are included to form a 'citation index'. For instance if I cited a paper in the Physical Review, volume D35, page 1001 it would be entered as PHRVA,D35,1001 and later searchers would be able to find all papers which cite PHRVA,D35,1001 including mine.
- 5. Extensive subject index terms are added by physicists at DESY to make very specialized topic searching possible.
- 6. A code is added to every conference paper making it possible easily to find many of the papers for a particular particle physics conference long before the proceedings is published.
- 7. If the preprint originated as an electronic preprint, as a majority now do, the e-print number is included and indexed.

THE IMPORTANCE OF RAPID COMMUNICATION

Because experiments are very expensive and have a very long lead time, particle physicists have for many years used preprints, the prepublication versions of both theoretical and experimental research results Journal publication has been very important as a stamp of approval but too slow to meet the day to day information requirements of this community. SPIRES-HEP, in fact, got its start from the need for a rapidly updated database which included preprints.

To this day, preprints and their kin are added immediately to the HEP database, mostly within 24 hours of arrival and there are no backlogs.

The other side of this is the ephemeral nature of preprints. That is, a preprint has in the past been only important until the day it was published in a journal.

Between 70 and 80% of all particle physics preprints which are recorded in the SLAC-HEP database are eventually published. The SLAC and DESY libraries track all of the major journals in particle physics, annotate the records in SLAC-HEP with the publication information and in the past have gladly discarded the dead preprints.

THE REVOLUTION OF 1991/92: E-PRINT ARCHIVES and the WORLD-WIDE-WEB

It had long been a source of discontent in the particle physics community that physicists in large laboratories had better access to the important preprint literature than those in small and/or distant departments in remote countries. The dream of having every paper available immediately to anyone who needed it had long been a subject of sometimes heated discussion but little action.

In 1991, Paul Ginsparg, a theoretical physicist now at Los Alamos National Laboratory, decided to take advantage of the fact that most of the younger physicists had done their theses using Donald Knuth's TeX formatting program. Using TeX, it was possible to transmit complex mathematical notation in a simple ASCII format across computer networks.

He and an associate devised a simple server (using his NeXT computer) which allowed physicists to post TeX 'source' in a certain format. A list of interested physicists would be notified each morning by e-mail of the new postings which they could then send for, again by e-mail. After that, it was up to them to produce a printable version of the TeX on their own computers. The TeX source was stored in ascii text files on an FTP server. Each paper was assigned a serial number based on year and month at the moment of posting, i.e. HEP-TH 9404001 would be the first April Fool's Day preprint posted to the HEP-TH server in 1994.

The scheme caught on almost immediately, soon there were more groups and more posting lists, hep-th, hep-ph, etc. etc. Into this rapidly evolving system came the contribution of the experimentalists. Tim Berners-Lee at CERN, the European Particle Physics Laboratory in Switzerland (a consortium of 12 nations) introduced the World-Wide-Web as a method for large groups of experimentalists to stay in touch and share information. We learned of it in 1991 when a SLAC physicist, Paul Kunz, returned from CERN with the news that WWW might be the perfect solution to our problem of how to make the SLAC-HEP database more accessible over the Internet. The result was the first Web server in the U.S. at SLAC serving the SPIRES-HEP database to the Internet world.

Since we already had to TeX all those e-prints from Ginsparg's servers, we soon took the next step which was to provide postscript versions with hypertext links to the SPIRES-HEP database. By now, we had learned to generate HTML (the hypertext markup language) directly from SPIRES, our database management system and it was easy to generate WWW URL's for the postscript versions. This soon became a very popular service since it meant that individual physicists no longer had to struggle with the vagaries of non-standard TeX.

The introduction of physicist Tony Johnson's MidasWWW, the first usable graphical Web browser (X-based) to handle compressed postscript, meant that all those Unix-based physicists out there could actually look up a paper in SPIRES-HEP and then see it. The playing field was suddenly level for particle physicists on the Internet and most of them by now had some kind of network access. Soon, the introduction of NCSA's Mosaic browser speeded up the acceptance of WWW as the network cruiser of choice. Usage statistics for SPIRES-HEP via WWW began to soar.

Addis-4/95



S985 - 1

· ·

CONVERTING E-PRINT TeX PAPERS TO VIEWABLE POSTSCRIPT

March 1995

- * Each night papers submitted to the e-print archives (acc-phys, astro-ph, gr-qc, hep-ex, hep-lat, hep-ph, hep-th, nucl-ex, nucl-th, and quant-ph) are downloaded at SLAC and automatically TeXed** to produce postscript output. Currently about 700 papers per month are received in this way.
- * Papers that fail the automatic TeX process (about 40%) are manually processed the next day at either the SLAC Library or by our colleagues at the DESY Library in Hamburg and posted to the SLAC postscript server.
- * All papers are tested for viewability on the Web and printability.
- * Authors are contacted by the SLAC library staff and asked to provide figures (either as Postscript or by fax). Faxed figures are converted to postscript at SLAC and posted as separate ps.Z files..
- * Input to the SPIRES-HEP citation index is also extracted from the TeX source during the automatic processing and stored for use by the inputters during the cataloging process.
- * Postscript documents are stored at SLAC and made available through FTP and HTTP (WWW).
- * Postscript documents are linked to bibliographic records in the SLAC SPIRES-HEP database.

RESULT - Using SPIRES-HEP via WWW, physicists can now not only look up papers in a consistently structured database, but also read and/or print many of these papers from their desktops.

** The automatic TeX system was developed by Paul Mende at Brown U.

Addis - 4/95

SOME CONSEQUENCES OF THE E-PRINT/WWW REVOLUTION ON DAILY OPERATIONS OF THE SLAC LIBRARY

1. A steady drop in preprint circulation occurred from 1991 (first year of e-prints) until in 1995, the library ceased shelving paper copies of e-prints.

Mo/Yr	Preprints Circulated
March 1991 -	783
March 1992	438
March 1993	356
March 1994	190
March 1995	73

This results in redirection of some staff effort and less pressure on shelf space. An old measure of library effectiveness bits the dust!

- 2. New higher volume postscript printer had to be purchased. Most physicists still want it on paper if they're seriously interested in a preprint.
- 3. More public X-terminals were and are needed. Not everyone has a high quality X-terminal on his/her desk for viewing compressed postscript.
- 4. It is now 'impossible' to lose an e-print. Users have immediate access to the papers they need (no recalls necessary) and the library conducts fewer raids on known packrats.
- 5. More staff time is spent checking quality of the posted papers and troubleshooting TeX, bad figures, etc. and corresponding with authors.
- 6. Input to the HEP database is somewhat eased by the automated extraction of citation lists from the TeX originals (or at least the citations are more accurately input). Still is not practical to extract bibliographic information from the e-prints due to lack of standards.
- 7. The Library dropped paper publication of its 25 year old preprint list (PPF) and started distributing it via listserv and WWW. When you don't put out paper lists, you have to work harder to ensure that everyone knows how to get your information and has access.
- 8. Preprint acquisition by mail is no longer a major effort. Instead, there is considerable effort expended in cruising various laboratory postscript servers to obtain newly posted papers, enter them in the HEP database with their correct URL's Not all papers are posted to the e-print archives and a growing trend seems to be for organizations to control their own postscript.
- 9. The library staff works in an environment which is changing even more rapidly than usual. Much retraining has to take place, mostly informally. Fortunately, they're flexible, smart and learning more about chaos theory every day.

Addis-4/95

HOPES FOR THE FUTURE

* LINKS TO JOURNAL ARTICLES
 • VIEWABLE BY SUBSCRIBERS

- * FULL-TEXT STANDARDS (FIGURES)
- * AUTHORS SUPPLY VIEWABLE PAPERS

SPIRES-HEP World-Wide-Web Information Sheet

April 1995

1. For more information about the WWW software, including a complete list of browsers and how to obtain them, telnet to:

telnet.w3.org (no password). You don't need anything more than a dumb terminal for this.

2. To link to the SLAC SPIRES home page from your WWW browser, open URL:

http://www.slac.stanford.edu/FIND/spires.html

- 3. For more information about the WWW SPIRES connection, e-mail to H. Galic (hep@slac.stanford.edu).
- 4. To link to the general SLAC home page, open:

http://www.slac.stanford.edu/FIND/slac.html

5. To be completely successful, you'll need a browser which can cope with compressed postscript (ps.Z) .The most popular such browser is NCSA's Mosaic for X. It can be obtained by anonymous FTP from:

ftp.ncsa.uiuc.edu in directory /Mosaic.

6. Or you might want to try the new version of MidasWWW (version 2.1), the browser pictured in these pages, available via FTP from:

freehep.scri.fsu.edu

in directory /freehep/networking_email_news/midaswww (includes source code as well as executables for aix, sun4, hpux, osf, sgi and VMS.)

MidasWWW is especially well suited for database index searching since the entire search statement is easily visible and the search area is always present and active at the bottom of the page no matter where the scroll bar has been pulled.

Addis, SLAC

SEARCH EXAMPLES

The following search examples illustrate a few of the features and power of the WWW/SPIRES-HEP combination.

Full-text, including figures and equations, is viewable for over 17000 preprints which have appeared on physics e-print archives or which are stored on postscript servers at various laboratories. The full text files (ps.Z) are linked to the SPIRES HEP database so that users may search by author, title, etc. and then view or print the entire preprint.

A warning: the 'look' of our WWW interface changes frequently as we try to improve the usability so examples may not always exactly match what you see on the screen.

- **Example 1** Search by author, leads to viewing a paper with color postscript figures
- Example 2 A search by title and date finds the two papers announcing the discovery of the 'top quark.' We explore those papers and follow a reference to and article in Nuclear Physics on-line.
- Example 3 Citation search

EXAMPLE 1

Using WWW and SPIRES to view the full text of a scientific paper by authors Hata and Langacker in the SPIRES-HEP database.

The paper could be found in various ways such as:

FIND AUTHOR HATA and LANGACKER FIND BULL hepph 9308252 FIND a hata, n and title updated msw

The document address is calculated from information stored in a SPIRES database called ABSTRACTS which is used for tracking e-print archive papers thru the TeXing and postscript creation process in the SLAC Library.

A SPIRES 'format' dynamically creates the HTML for WWW. An example is shown here.

This example uses the MidasWWW browser running under X-windows and views a postscript version of the paper with color postscript figures.

File Postscript Navigate Customize Documents Manuals

Document: [http://www.slac.stanford.edu/FIND/spires.html

Stanford Public Information REtrieval System SLAC Last update: 21 Mar 1995 Welcome to SLAC-SPIRES Information Retrieval System A variety of SLAC's databases of interest to high-energy physics community is now made available via WWW. This project is still in the experimental phase and we do appreciate your feedback. If a caretaker of a particular database is not listed, please address your comments to: library@slac.stanford.edu The following SLAC's information sources are currently accessible to WWW users: Books: SLAC Library book catalog. lick here **Preprints**: HEP preprint database. Contains bibliographic summaries of more than 300,000 particle physics papers. Included are preprints, journal articles, technical reports, thesis, etc. Searchable by author, title, report number, institution, collaboration, and more. Find citations of your favorite author or article. View postscript versions of selected preprints, read abstracts of e-print archive papers. Neepingre help? Try also the latest version of our (still) highly experimental forms search' in the HEP database. Send comments/suggestions related to the forms to: tony_johnson@slac.stanford.edu **Recent e-Prints**: Useful in searching for recent high-energy physics e-prints ("bulletin-board" papers) not yet covered by the HEP database. Find abstracts and viewable postscript (made at SLAC) of articles posted today, yesterday, in the last seven days, week before that, or anytime. Preferred access to older e-prints is through the HEP preprint database (above). Hepnames: World-wide e-mail directory of people related to particle physics. Includes SLAC physicists, personnel. Keyword:

Search...

Go Back Previous

Next

Save...

Clone

Close Window



Ŵ

Title: FIND AUTHOR HATA AND DATE AFTER 1990 (HEP-SPIRES-SLAC)	
locument: http://slacvm.slac.stanford.edu.5080/find/spiform?author-Hata&title-&affiliation-	
Format: HTML	
Show Document Source Search] Help Cancel	File Posiscript Navigate Customize Documents Manuals Her
) b> THE UPDATED HSW ANALYSIS AND THE STANDARD SOLAR HODEL UNCERTAINTIE	Document: http://lbnext.slac.stanford.edu:5080/hep-ph/9308/9308252.ps.Z
aoya Bata, Paul Langacker (Penn U.). UP2-0581-Y, Aug 1993. 15pp. U talk given at International Workshop on Supersymmetry and Unific	This is a multipage Postscript document, select page:
mdamental Interactions (SUSY 93), Boston, MA, 29 Mar - 1 Apr 1993	Page 2 - +ITLE PAGE
In "Boston 1993, Proceedings, Supersymmetry and unification of amental interactions, SUSY 93" 562-572. and Pennsylvania Univ.	• Fago 2 • Pago 3
delphia - UP2-0581T (93/08, rec. Aug.) 15 p. and (314708). -Print Archive: hep-ph/9308252	• Fago 2 • Pago 4
	• Page 5
dd×a href=hep?fin+key+2797020+(using+wwwrefs> Deferences	• Page 6
l> Abstract_and_Paper <td>• Page 7</td>	• Page 7
LLC postscript depository d>Link to Low Alal	• Page 8
rver	 Pago 6 Pago 7 Pago 8 Pago 9 Pago 10 Pago 11 Pago 12 Pago 13 Pago 14 Pago 15
<pre> Conference_Info <)</pre>	• Page 10
	• Page 11
CORY OF THEORIES' APPROACH TO STRING THEORY.	· Page 12
oyuki Hata (Kyoto U.). KUNS-1212, Aug 1993. 19pp.	· Page 13
blished in <1>Phys.Rev.D50:4079-4087,1994. 1	• Page 14
<pre>ka href=hep?fin+key+2791102+(using+wwwrefs> References </pre>	• Page 15 Page 15
 Keywords 	
<pre>xa href=hep?fin+c+PHEVA,D50,4079> Citation_Search </pre>	Note
AC postscript depository >Link to Los Ala	Once in the document you can use the Noxt and Provious buttons to
veink to Log klai	go to adjacent pages, and the Go Back button to return to this index.
TROPHYSICAL SOLUTIONS ARE INCOMPATIBLE NITH THE SOLAR NEUTRINO D	
Budman, N. Hata, P. Langacker (Penn U.). UPR-0572-7, Jun 1993.	Keyword:
ublished in <i>Phys.Rev.D49:3622-3625,1994.</i> (Hata first autho)	
Print Archive: bep-ph/9306212	Go Back Previous Next Save Search Clone
<pre>>> href=hep?fin+key+2757478+(using+wwwrefs> References </pre>	
d> Keywords d> Citation_Search 	
<pre>> Abstract_and_Paper </pre>	
stscript depository	

Help





•



.

EXAMPLE 2:

We look for the title words 'TOP QUARK OBSERVATION' and find the two March 1995 papers from Fermilab announcing the important discovery of the 'top quark.'

We examine the reference list for one of these papers and follow a reference to a paper published in Nuclear Physics which is available from the Elsevier WWW server to Nuclear Physics subscribers.







Ŵ

The hot linked with File Postscript Navigate Customize Documents Manuals File Postscript Navigate Customize Documents Manuals Help Help Title: FIND SPICITE NUPHA B403,633 (HEP-SPIRES-SLAC) Title: Untitled (http://siacvm.siac.stanford.edu:5080/find/hep?fin+key+3100952+(using+wwwrefs) Document: [http://slacvm.slac.stanford.edu:5080/find/hep?fin+key+3100952+(using+wwwrefs Document: http://slacvm.slac.stanford.edu:5080/find/hep?fin+spicite+NUPHA_B403,833 Some of the references, mostly to bulletin-boards and published journal articles, from Database: HEP (SPIRES-SLAC) the paper: Observation of the top quark (Only the first author is displayed, where Search Command: FIND SPICITE NUPHA, B403, 633 known) Result: 1 document found: HIGHER ORDER CORRECTIONS TO JET CROSS-SECTIONS IN HADRON Phys. Rov. Lott. 72, 2138 (Abachi: Search For The Top Quark In P Anti-P Colli...) Phys. Bov. D50, 2966 (Abe: Evidence For Top Quark Production In Anti-P P ...) COLLIDERS. By W.T. Giele (Fermilab), E.W.N. Glover (Durham U.), David A. Kosower (CERN Phys. Bov. Lott. 73, 225 (Abe: Evidence For Top Quark Production In Anti-P P ...) &Saclay), FERMILAB-PUB-92-230-T, (Received Jan 1993). 36pp. Phys.Lett. B321, 254 (Laenen: Top Quark Production Cross-Section ...) Nucl.Instrum, Meth. A338, 185 (Abachi: The D0 Detector ...) Published in Nucl. Phys. B403:633-670,1993. e-Print Archive: hep-ph/9302225 Nucl. Phys. B403, 633 (Giele: Higher Order Corrections To Jet Cross-Secti...) Comput. Phys. Commun. 67, 465 (Marchesini: Herwig: A Monte Carlo Event References Generator ...) Keywords Clion here to expand this reforence Citation_Search Abstract and ToX Source. Link to Los Alamoa server Postscript_Version from CERN Source from Nucl. Phys. server (Access may be restricted) Link to the main HEP (SPIRES-SLAC) page Dies anvaigner mint. public Keyword: Keyword: Go Back Previous Next Save ... Search ... Clone Close Window Go Back Previous Next · Save ... Search ... Clone Close Window



EXAMPLE 3:

We look for author Paul Ginsparg and use the special format to do citation searching on his publications.

¥

File Postscript Navigate Customize Documents Manuals Help	File Postscript Navigate Customize Documents Manuals
The HEP Preprint database	SLAC. 3 Apr 1995
The HEP preprint database contains bibliographic summaries of more than 280,000 particle physics papers. Included are preprints, journal articles, technical reports, thesis, etc. Need help? Choose ? below for help on any field. Search Parameters	SLAC.3 Apr 1995 The SPIRES command is: FIND AUTHOR ginsparg AND DATE MTER 1989 (In HEP USING WWW Change Format to Default & Full & Citation
Author:_ ginspard ?	Result:
?	Database: HEP (SPIRES-SLAC) Search Command: FIND AUTHOR GINSPARG AND DATE AFTER 1989 Result: 9 documents found:
Collaboration: Aleph Other [1) LECTURES ON 2D GRAVITY AND 2D STRING THEORY. By Paul Ginsparg (Los Alamos), Gregory Moore (Yale U.). YCTP-P23-92, (Received Apr 1993). 230pp. In Kyoto U., Yukawa Inst. Library only. Lectures given at TASI summer school, Boulder CO, June 11-19, 1992. Yale Univ. New Haven - YCTP-P23-92 (92,rec.Apr.93) 197 p. Los Alamos Nat. Lab LA-UR-92-3479 (92,rec.Apr.93) 197 p. e: LANL hep-th/9304011, In *Boulder 1992, Proceedings, Recent directions in particle theory* 277-469. and Yale Univ. New Haven - YCTP-P23-92 (92,rec.Apr.93) 197 p. and Los Alamos Nat. Lab LA-UR-92-3479 (92,rec.Apr.93) 197 p.
Result format	e-Print Archive: hop-th/9304011 References
Show only number of matches	Ney Worse
Show all matches using Default format <u>1</u> . Perform Search Clear Form	Cliation Search Contraction of the Search Characteristic Search Contractions of the Search Characteristic Search Contractions of the Search Characteristic
Keyword: Go Back Previous Next · Save Search · Clone Close Window	Keyword: Go Back Previous Next · Save Search · Clone Close Window

