The World Wide Web Virtual Library

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Knowledge Dissemination
The WWW Virtual Library

Michael Chapman . . . . . Arthur Secret
Raoul Sanchez . . . . . Lisa Micah

Paper presented to *THE INTERNET SOCIETY*
Development Special Interest Group
Wednesday October 18th, 2006 — CERN, Geneva

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

1. History
2. What we are
3. The VL in 2006
4. The Future
5. Conclusion

Historical Introduction

Ethos

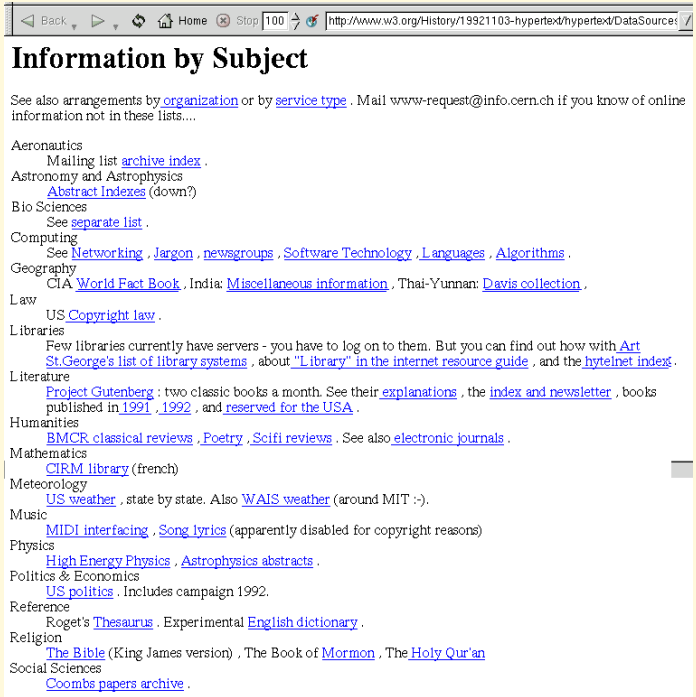
The WWWVL in 2006

The WWWVL and the . . .

Conclusion

1. Historical Introduction

Overview.html



Back Home Stop 100 http://www.w3.org/History/19921103-hypertext/hypertext/DataSources

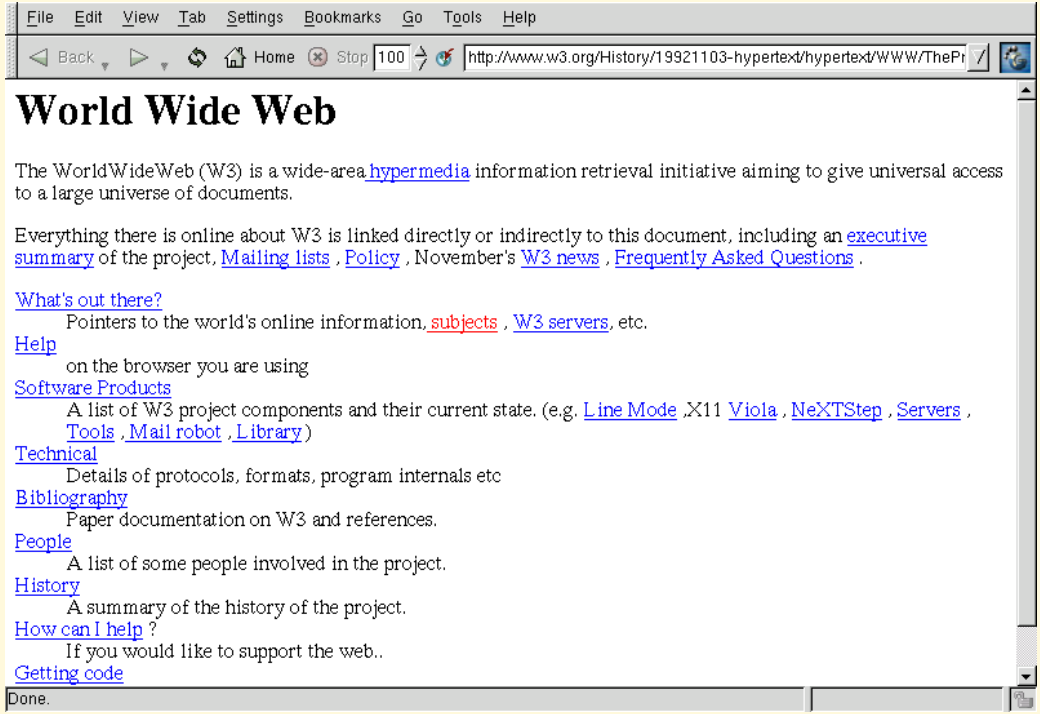
Information by Subject

See also arrangements by organization or by service type . Mail www-request@info.cern.ch if you know of online information not in these lists....

Aeronautics
Mailing list archive index .
Astronomy and Astrophysics
Abstract Indexes (down?)
Bio Sciences
See separate list .
Computing
See Networking , Jargon , newsgroups , Software Technology , Languages , Algorithms .
Geography
CIA World Fact Book , India: Miscellaneous information , Thai-Yunnan: Davis collection .
Law
US Copyright law .
Libraries
Few libraries currently have servers - you have to log on to them. But you can find out how with Art St.George's list of library systems , about "Library" in the internet resource guide , and the hytelnet index .
Literature
Project Gutenberg ; two classic books a month. See their explanations , the index and newsletter , books published in 1991 , 1992 , and reserved for the USA .
Humanities
BMCR classical reviews , Poetry , Scifi reviews . See also electronic journals .
Mathematics
CIRM library (french)
Meteorology
US weather , state by state. Also WAIS weather (around MIT :-).
Music
MIDI interfacing , Song lyrics (apparently disabled for copyright reasons)
Physics
High Energy Physics , Astrophysics abstracts .
Politics & Economics
US politics . Includes campaign 1992.
Reference
Roget's Thesaurus . Experimental English dictionary .
Religion
The Bible (King James version) , The Book of Mormon , The Holy Qur'an
Social Sciences
Coombs papers archive .

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

November 3rd 1992:



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

"… there would be the occasional 'Hey, I've just set up a server, and it's dead cool. Here's the address.'

With each new message I would enter in info.cern.ch a hyper-text link to the new web site …"

"In July and August 1991 there were from ten to one hundred 'hits' (pages viewed) a day."

Tim Berners-Lee *Weaving the Web* (Harper Business ISBN 0-06-251587-X), pp. 47 & 49.

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the…

Conclusion

Information by Subject

See also arrangements by organization or by service type . Mail www-request@info.cern.ch if you know of online information not in these lists....

Aeronautics
Mailing list archive index .
Astronomy and Astrophysics
Abstract Indexes (down?)
Bio Sciences
See separate list .
Computing
See Networking , Jargon , newsgroups , Software Technology , Languages , Algorithms .
Geography
CIA World Fact Book , India: Miscellaneous information , Thai-Yunnan: Davis collection ,
Law
US Copyright law .
Libraries
Few libraries currently have servers - you have to log on to them. But you can find out how with Art St.George's list of library systems , about "Library" in the internet resource guide , and the hytelnet index .
Literature
Project Gutenberg : two classic books a month. See their explanations , the index and newsletter , books published in 1991 , 1992 , and reserved for the USA .
Humanities
BMCR classical reviews , Poetry , Scifi reviews . See also electronic journals .
Mathematics
CIRM library (french)
Meteorology
US weather , state by state. Also WAIS weather (around MIT :-).
Music
MIDI interfacing , Song lyrics (apparently disabled for copyright reasons)
Physics
High Energy Physics , Astrophysics abstracts .
Politics & Economics
US politics . Includes campaign 1992.
Reference
Roget's Thesaurus . Experimental English dictionary .
Religion
The Bible (King James version) , The Book of Mormon , The Holy Qur'an
Social Sciences
Coombs papers archive .

Coombs Papers

This archive is a public archive of academic papers in the social sciences, and also in Pacific Rim geography and religious studies.

Chose your nearest site

Main archive site is at Australian Nat'l. University, Canberra with daily copies to Washington University, St. Louis, USA and with weekly copies to the US national UUCP gateway, California, USA and Univ. of Nth Carolina, USA .

They are just about to release two WAIS servers - one with inf. about Pacific Manuscripts microfilms collection at ANU, the other one about the anthropological/sociological research in the Thai-Yunnan arrea of the SE Asia. The launching is going to happen tomorrow, 21st May 92.

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

August 1992: There are about 20 Web servers in existence.

Ciolek: Global Networking: a Timeline.

VL (with logo) first known screenshot 1993

(by October 1993 — 200+ servers . . .)

Ciolek: Global Networking: a Timeline.

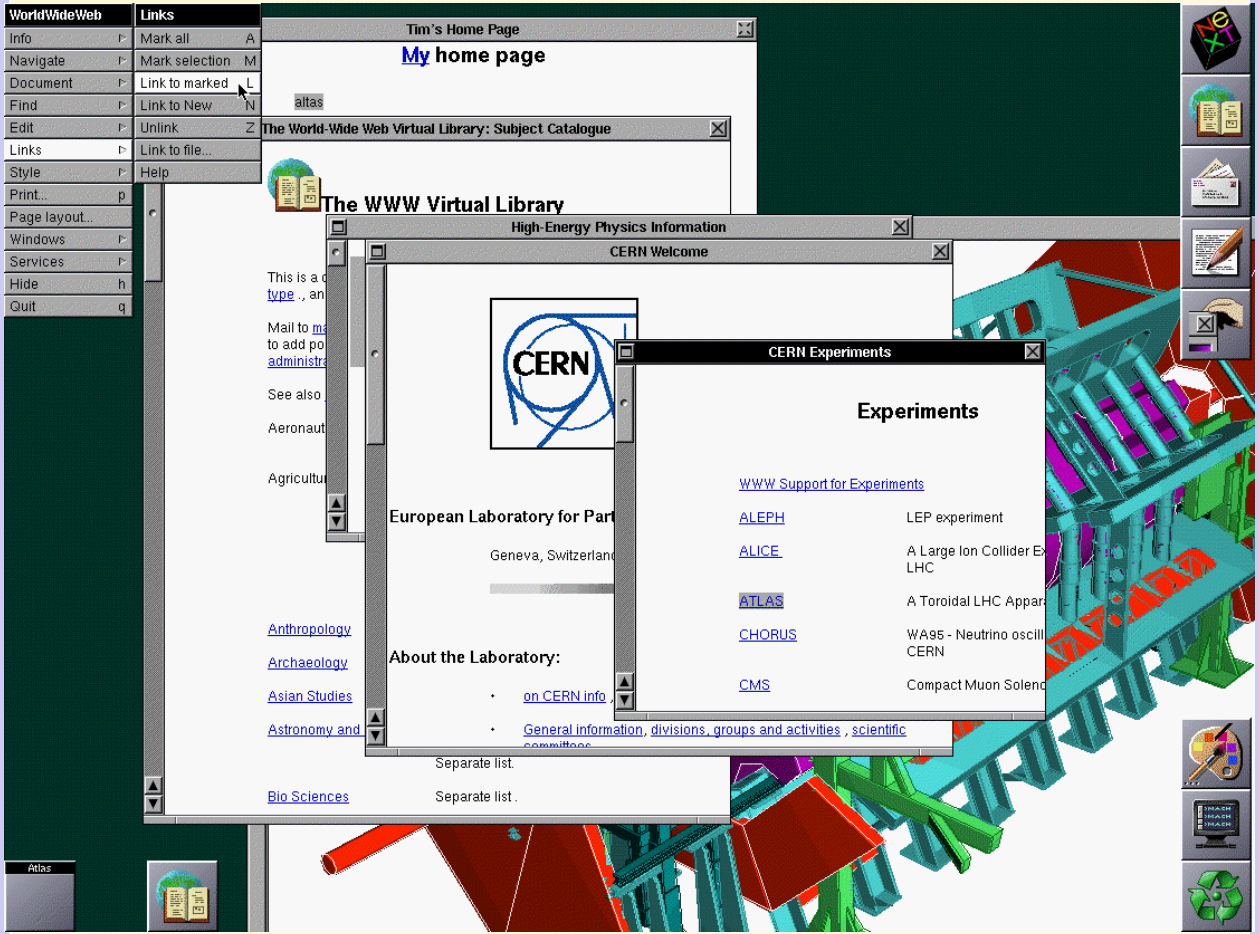
Historical Introduction

Ethos

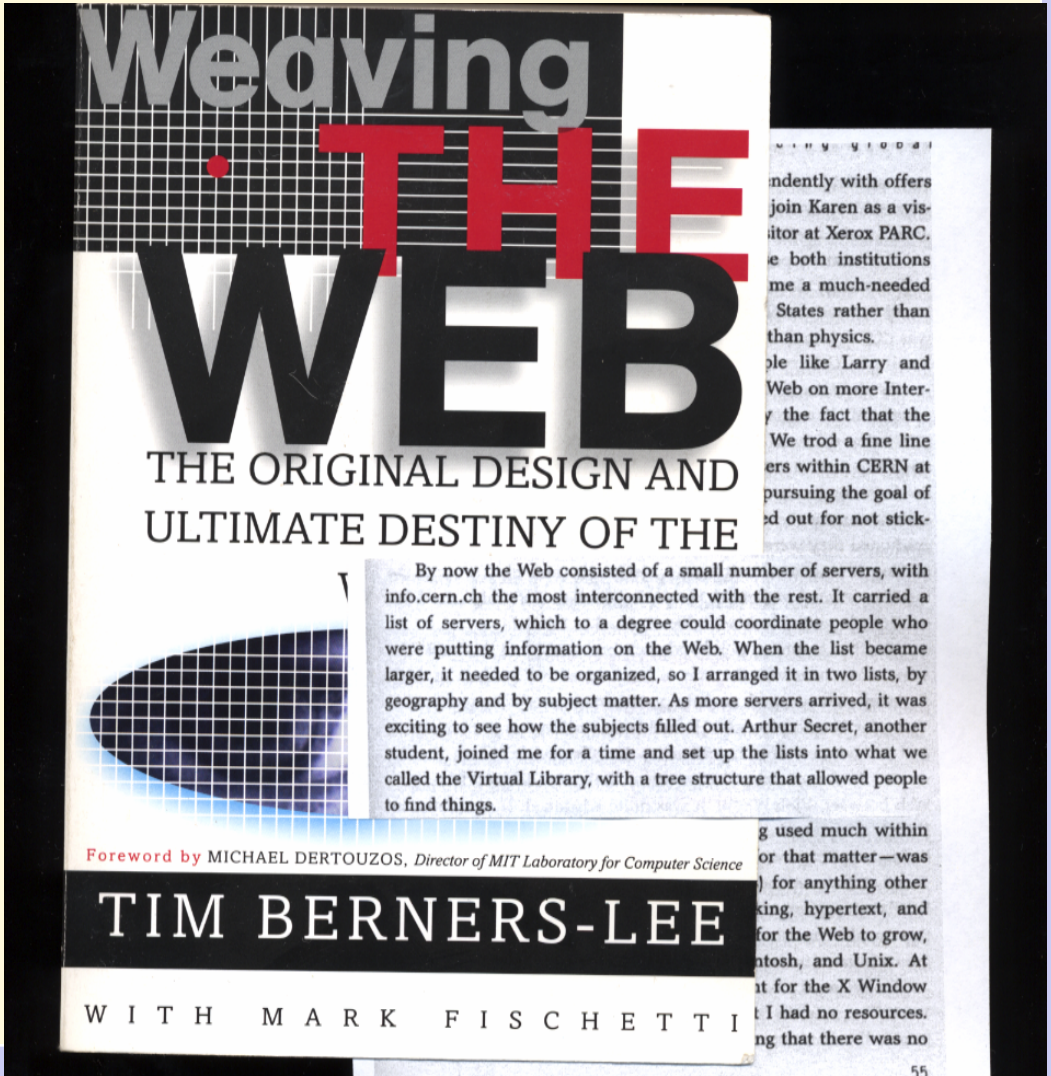
The WWWVL in 2006

The WWWVL and the . . .

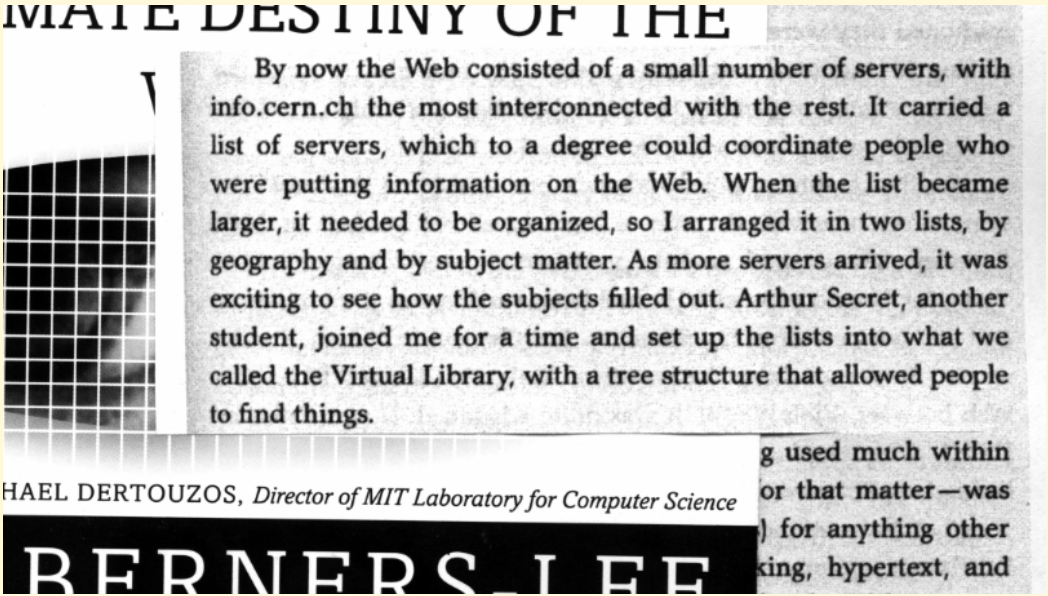
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



MATE DESTINY OF THE

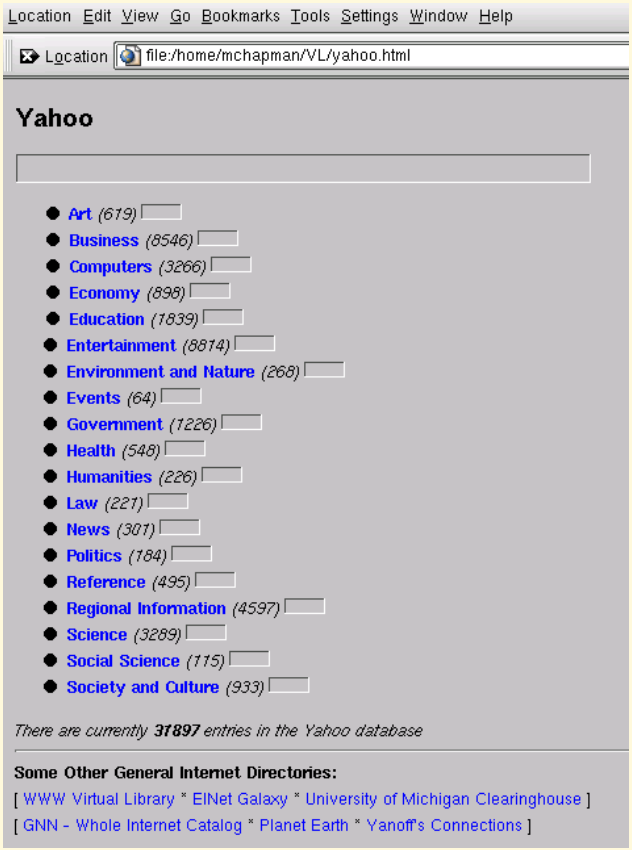
By now the Web consisted of a small number of servers, with info.cern.ch the most interconnected with the rest. It carried a list of servers, which to a degree could coordinate people who were putting information on the Web. When the list became larger, it needed to be organized, so I arranged it in two lists, by geography and by subject matter. As more servers arrived, it was exciting to see how the subjects filled out. Arthur Secret, another student, joined me for a time and set up the lists into what we called the Virtual Library, with a tree structure that allowed people to find things.

g used much within
or that matter—was
) for anything other
king, hypertext, and

HAEL DERTOUZOS, *Director of MIT Laboratory for Computer Science*

BERNERS-LEE

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

March 1995/2005:

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

WWWVL

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

2. Ethos

- What we are
- What we are not
- Summary

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

2.1. What we are

The same concept from inception until now.

Like the Web itself, the VL is simple and logical.

Like the Web itself, it also has at its foundation the free distribution of free information.

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

2.2. What we are not

"It is futile to discuss whether the Index *or the table of* Contents *is the more useful in a book."*

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

Not a search engine . . .

Google no longer say, on their main welcome page, how many pages they index . . . but a search for “html” earlier this year produced 12,920,000,000 pages.

Compared with that 11 figure total the VL only refers to a *six* figure number of links/URLs.

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

Not an encyclopædic directory project . . .

Compared with **dmoz**/ODP, both volunteer:

"Just as the Oxford English Dictionary became the definitive word on words through the efforts of a volunteers, the Open Directory follows in its footsteps to become the definitive catalog of the Web."

but attributable *v.* "The Republic of the Web".

VL is more monographic as against dmoz's ambition to be encyclopædic.
(One could say 'academic' *v* general, but that has too many pejorative overtones.)

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

Not an encyclopædic directory project . . .

VL is attributable. VL is selective.

VL lacks a Wiki's speed of response.

Again 'different animals'.

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

Quality rather than breadth.

Indeed the one thing we do well is ‘focussed’ VLs.

We are determined to keep the original VL concept, but seeking to finance that without commercialisation will be the challenge.

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

2.3. Summary

The best paper comparison is perhaps with a small academic publisher.

We select / are selected by individuals who have a discrete collection of knowledge they wish to put before a (perhaps specialist) public. They contribute their knowledge, their expertise. We ‘market’ their œuvre, both actually by placing in the Catalog, but also by adding our ‘imprimatur’ to the work.

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

A collection of

* individual,

* attributable,

* autonomous,

* expert subject indices.

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

A collection of

* individual,

* attributable,

* autonomous,

* expert subject indices.

“It is futile to discuss whether the Index *or the table of* Contents *is the more useful in a book.”*

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

vlib.org

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

3. The WWWVL in 2006

- The Association
- Central Services
  1. Catalog
  2. VLsearch
  3. Topic Map
- Infrastructure

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

3.1. The Association

The WWWVL consists of over two hundred individual VLs, each maintained by individuals and institutions based throughout the World.

Many of those VLs are leading resources in their own fields of study, irrespective of their WWWVL membership.

VLs in English, French, German.
Parts of VLs in many, many languages.

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Services to members:

- link checking (?spell checking, accessibility checking, . . .)
- site establishment/design guidance
- discussion lists
- . . . meetings (first regional meeting, in Geneva this year)

Also . . .

- Press relations
- Copyright permissions
- ‘Outreach’ . . . today!

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

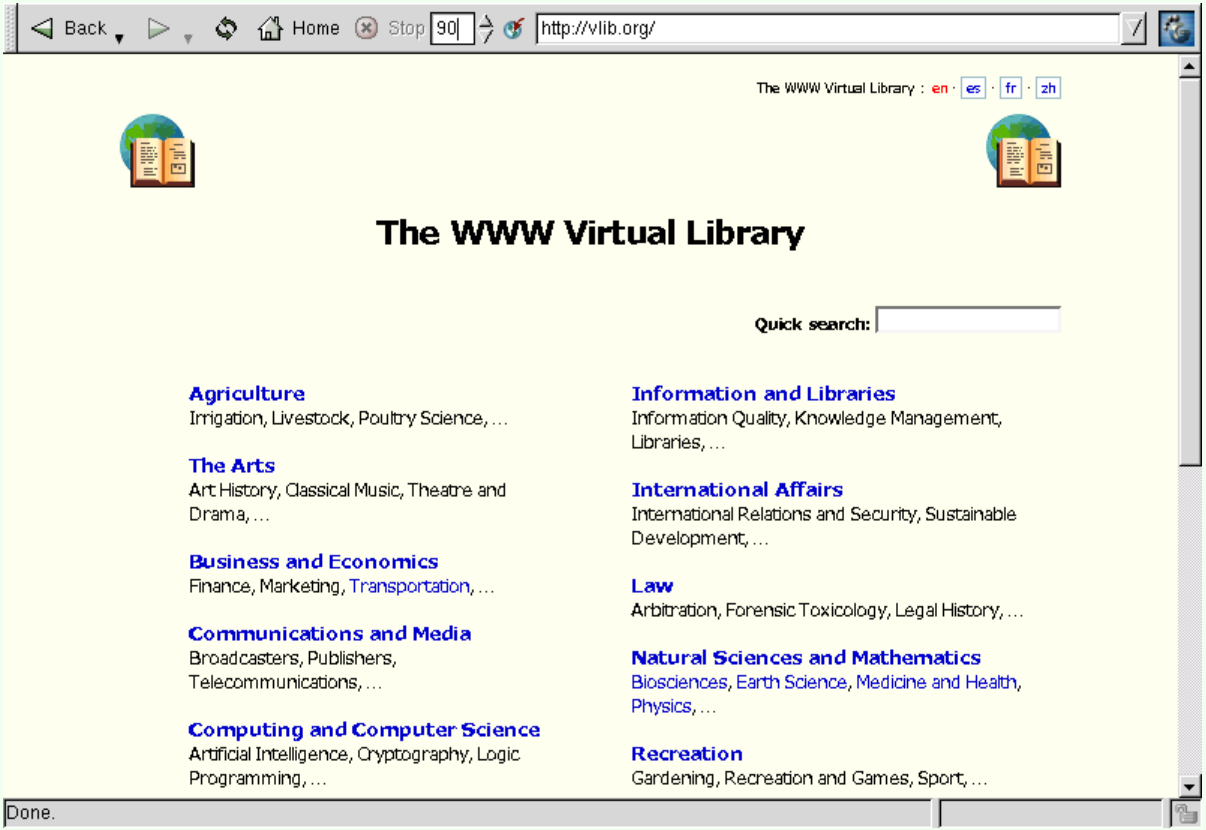
3.2. Central Services

The Web interface to/for the public.

- a hierarchical catalogue
- a search engine
- a topic map (. . . to be queriable)

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

3.2.1. Central Catalog

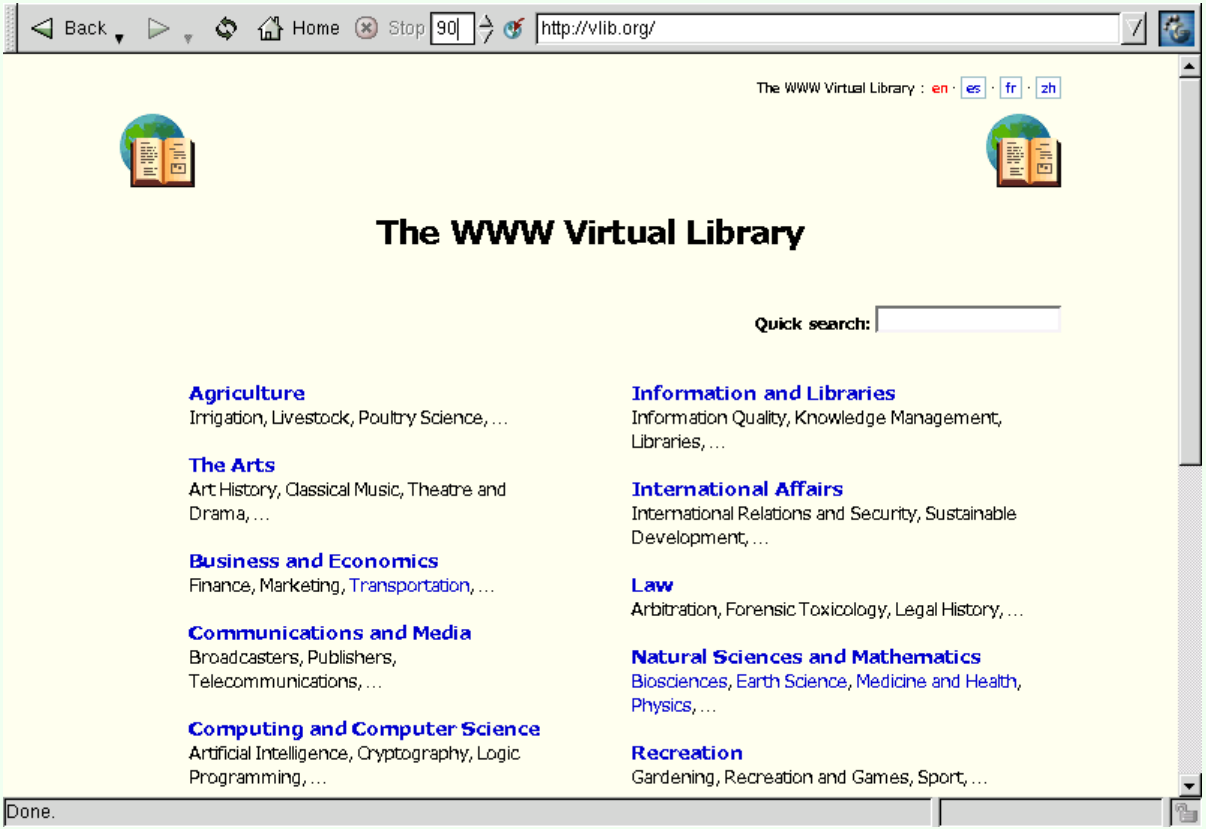


Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the...
Conclusion

Interfaces in various languages.

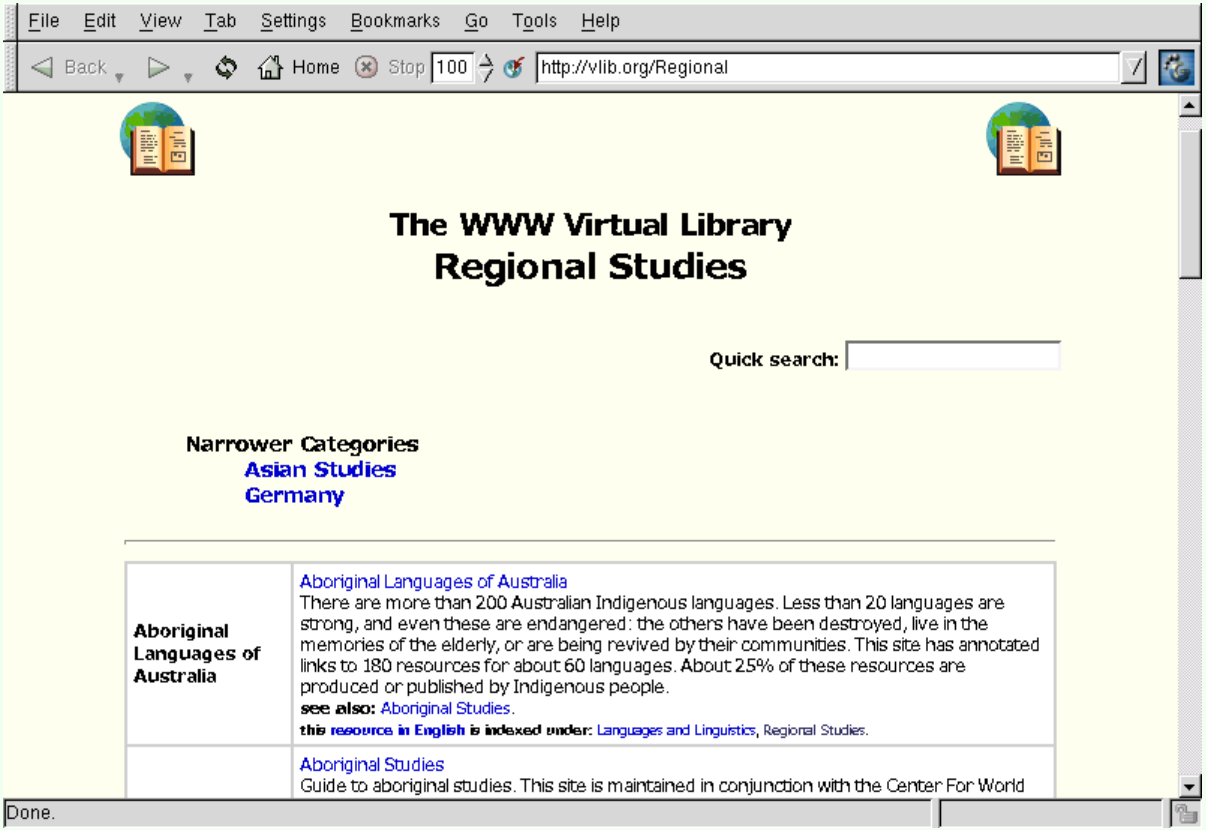
Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Looking for *Myanmar*

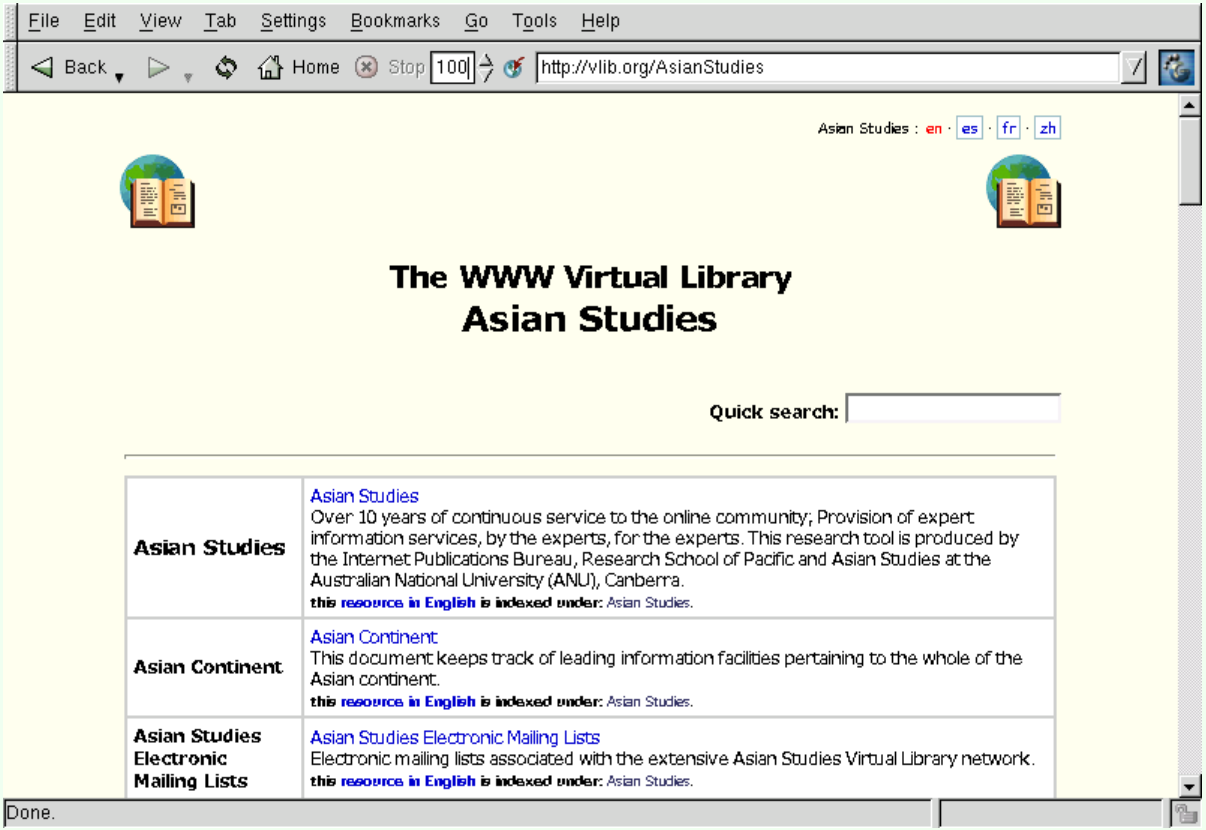


Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

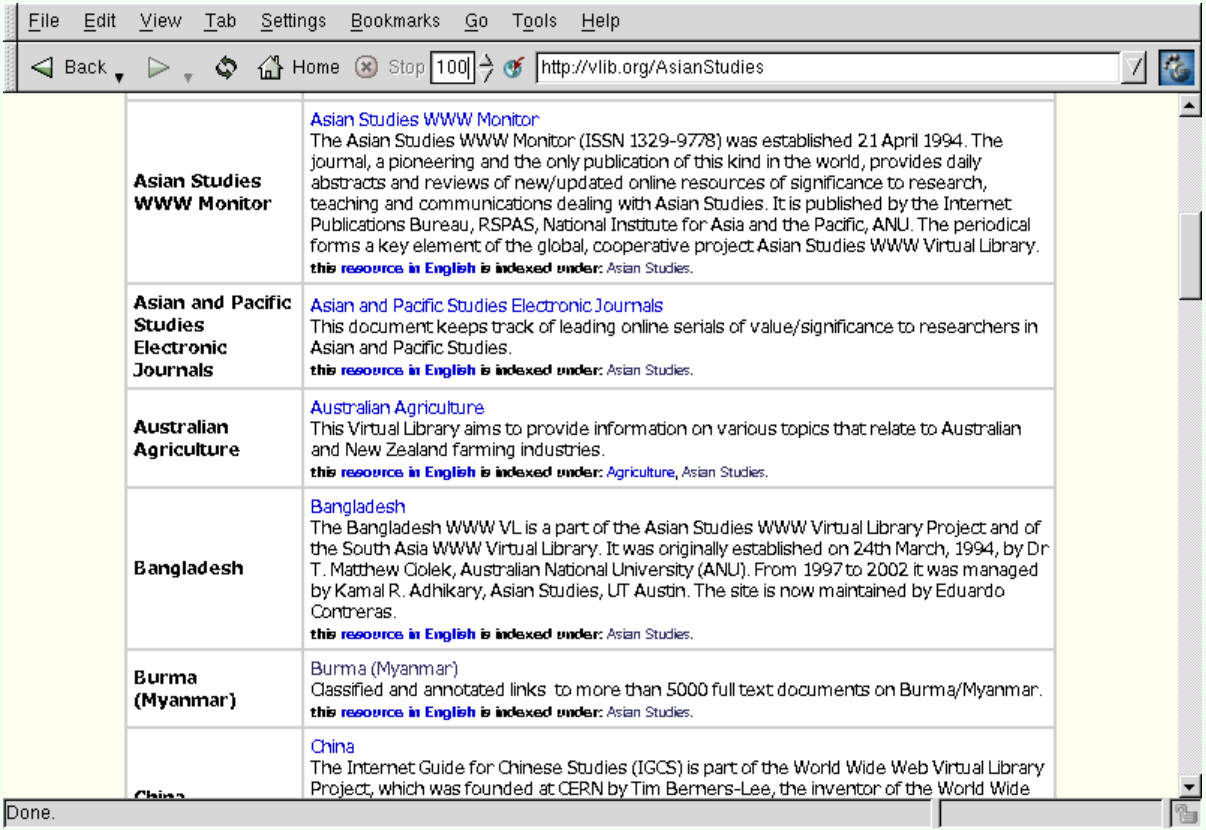
Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



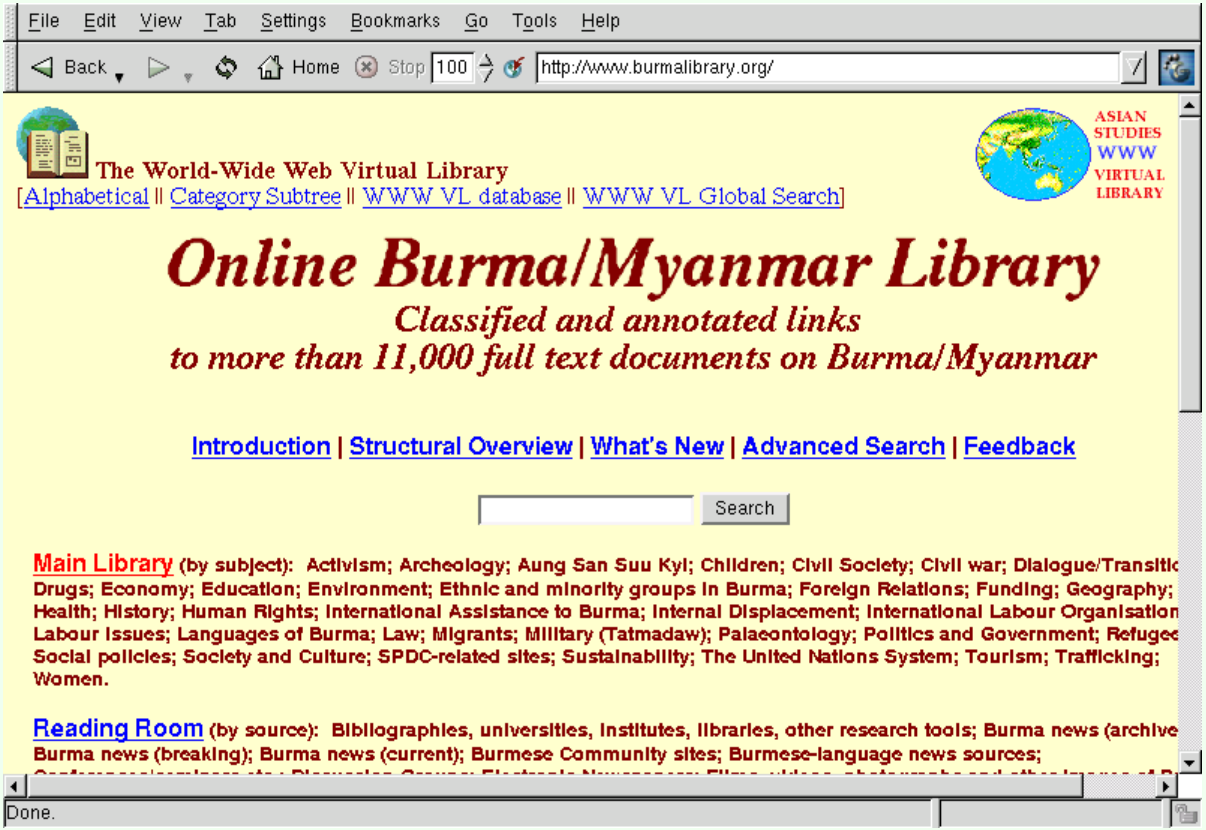
Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Usage:

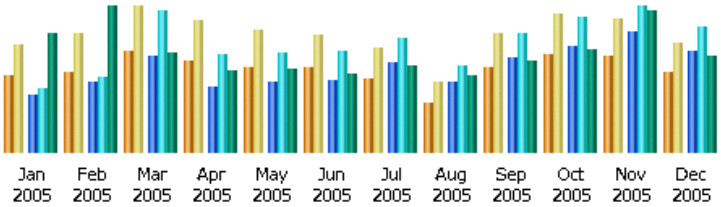
Summary

Reported period	Year 2005
First visit	01 Jan 2005 - 00:00
Last visit	31 Dec 2005 - 23:58

	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Viewed traffic *	**<= 879441** Exact value not available in 'Year' view	**1244901** (1.41 visits/visitor)	**3815042** (3.06 pages/visit)	**4911222** (3.94 hits/visit)	**15.46 GB** (13.02 KB/visit)
Not viewed traffic *			**3695561**	**4474906**	**4.69 GB**

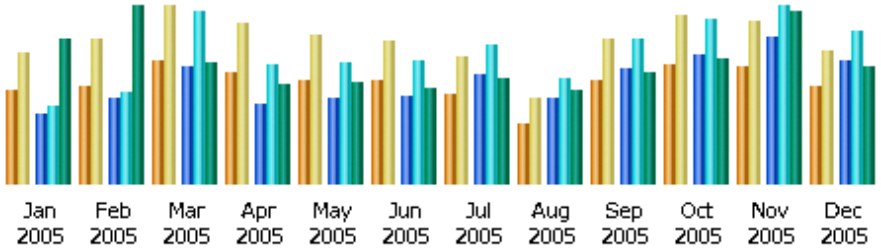
* Not viewed traffic includes traffic generated by robots, worms, or replies with special HTTP status codes.

Monthly history



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Monthly history



Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2005	67001	94207	217869	237972	1.52 GB
Feb 2005	70994	104547	263158	281958	1.88 GB
Mar 2005	88623	127805	359994	532927	1.29 GB
Apr 2005	80491	115451	247323	366342	1.06 GB
May 2005	74714	107258	267551	377381	1.07 GB
Jun 2005	73997	103155	270429	379142	1.00 GB
Jul 2005	64624	91394	335817	427027	1.11 GB
Aug 2005	43872	61664	263555	328160	1005.86 MB
Sep 2005	73973	104360	353982	447580	1.18 GB
Oct 2005	86546	121786	402057	509280	1.31 GB
Nov 2005	84576	116920	451640	548911	1.82 GB
Dec 2005	70030	96354	381667	474542	1.24 GB
Total	879441	1244901	3815042	4911222	15.46 GB

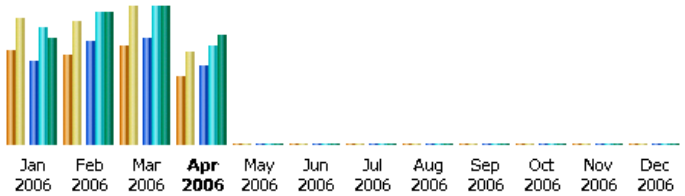
Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Summary

Reported period	Month Apr 2006				
First visit	01 Apr 2006 - 00:00				
Last visit	24 Apr 2006 - 02:01				
	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Viewed traffic *	**58879**	**79947** (1.35 visits/visitor)	**223431** (2.79 pages/visit)	**278443** (3.48 hits/visit)	**1.08 GB** (14.13 KB/visit)
Not viewed traffic *			**432011**	**451915**	**675.36 MB**

* Not viewed traffic includes traffic generated by robots, worms, or replies with special HTTP status codes.

Monthly history



Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2006	81070	109750	233339	328307	1.05 GB
Feb 2006	77888	107144	292084	370901	1.31 GB
Mar 2006	85490	119140	300524	386862	1.36 GB
Apr 2006	58879	79947	223431	278443	1.08 GB
May 2006	0	0	0	0	0

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Hierarchical catalogue, the ‘Catalog’:

- built from a database (actually a topic map), this allows:
  - rich cross–linking
  - entries to easily appear on more than one page
  - rapid, easy updating
- the most popular entry point for the VL, with approaching one million unique visitors per year.

Historical Introduction

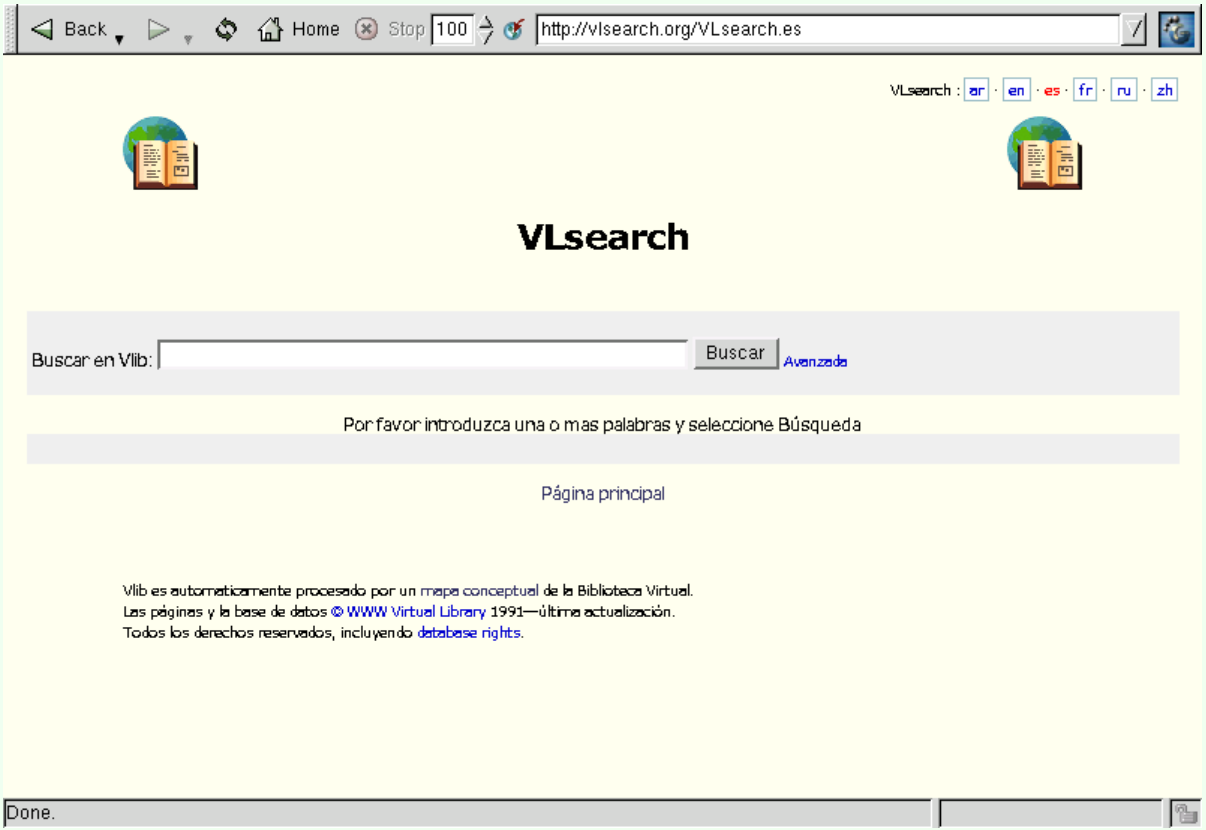
Ethos

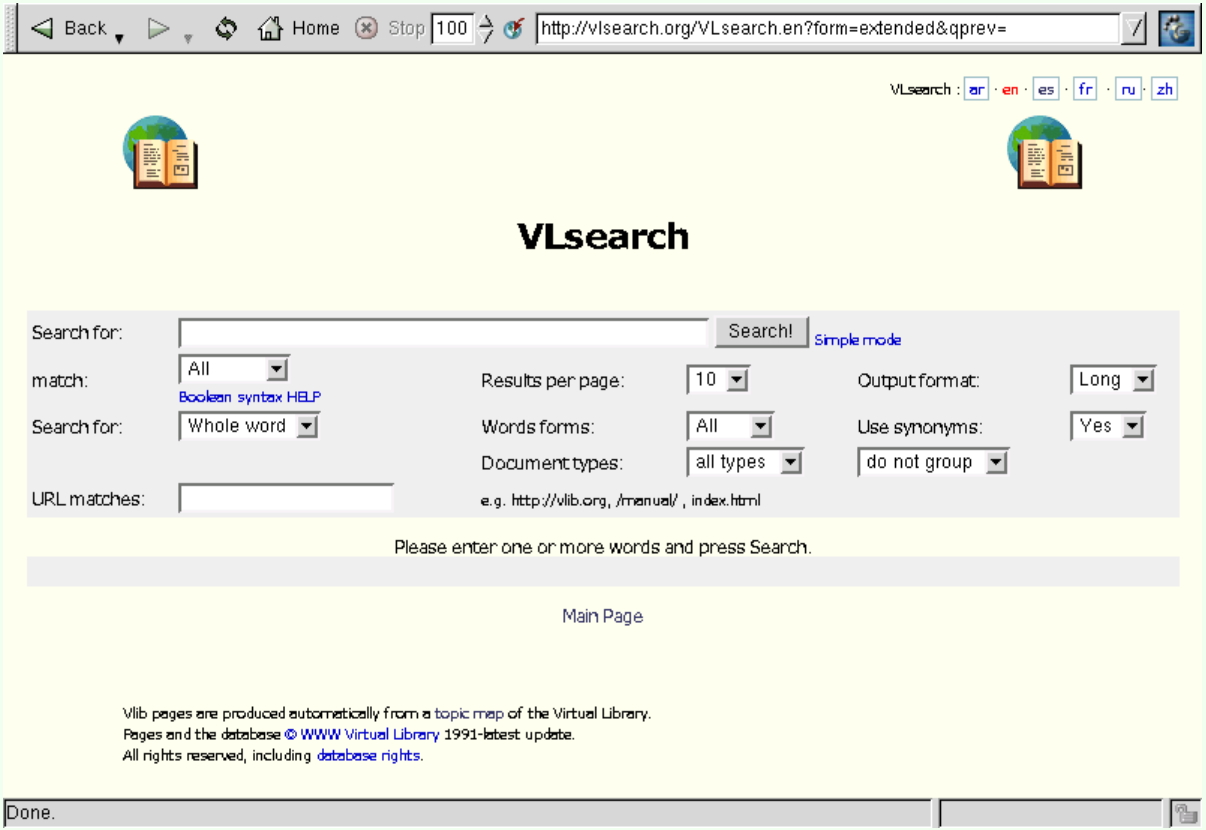
The WWWVL in 2006

The WWWVL and the . . .

Conclusion

3.2.2. VLsearch





Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Historical Introduction

Ethos

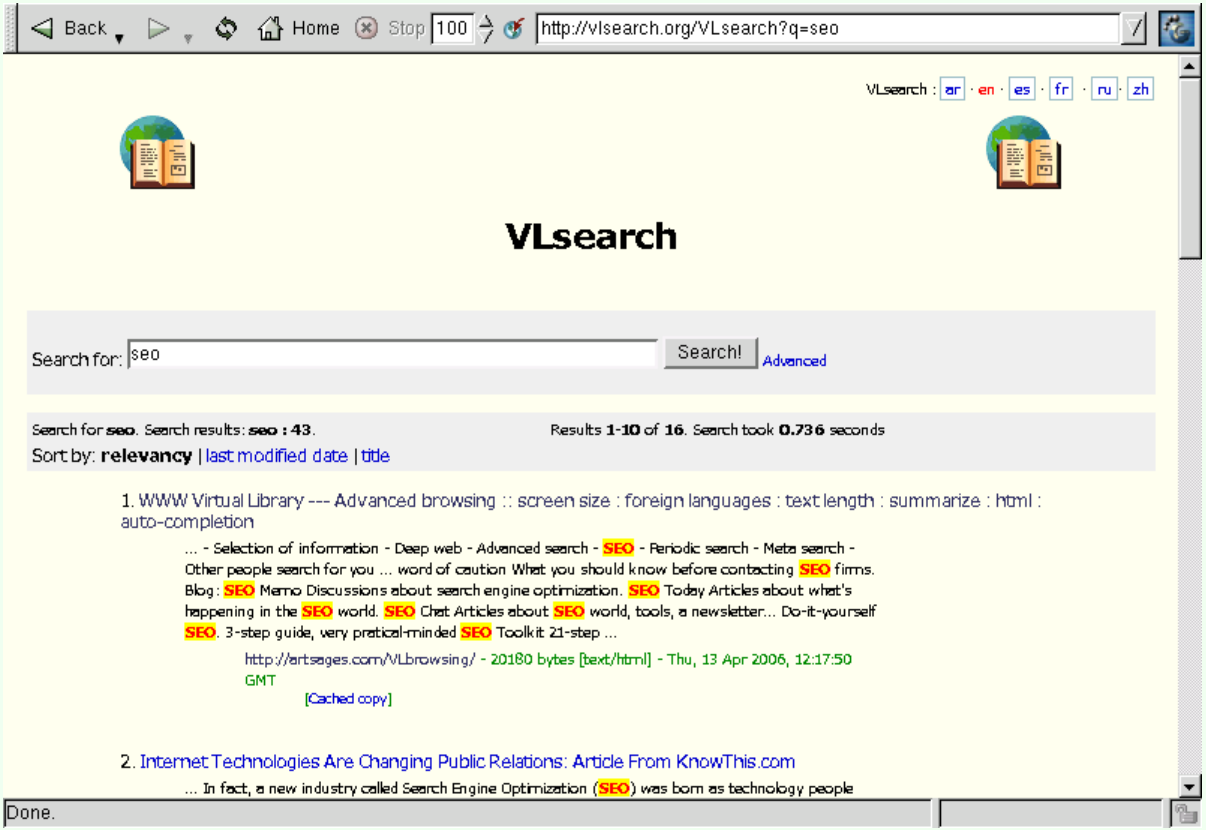
The WWWVL in 2006

The WWWVL and the . . .

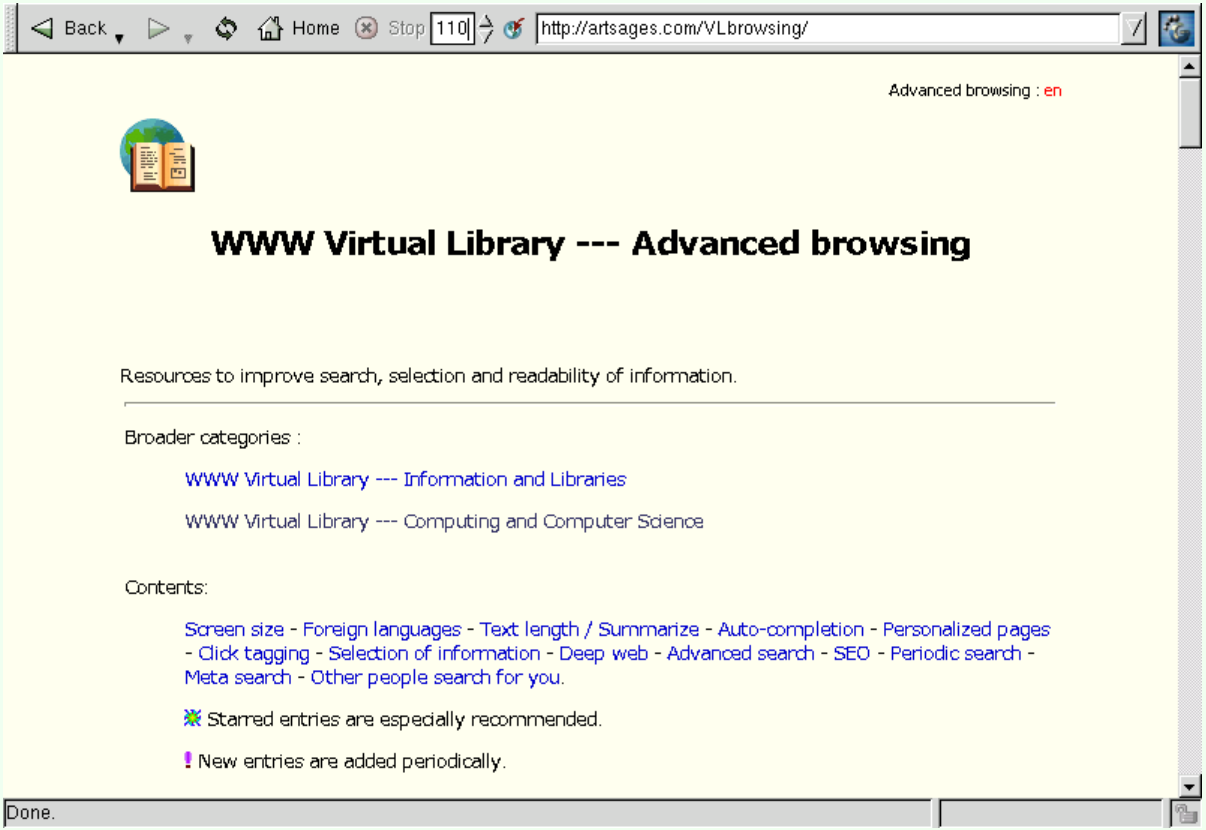
Conclusion

About *search engine optimisation*

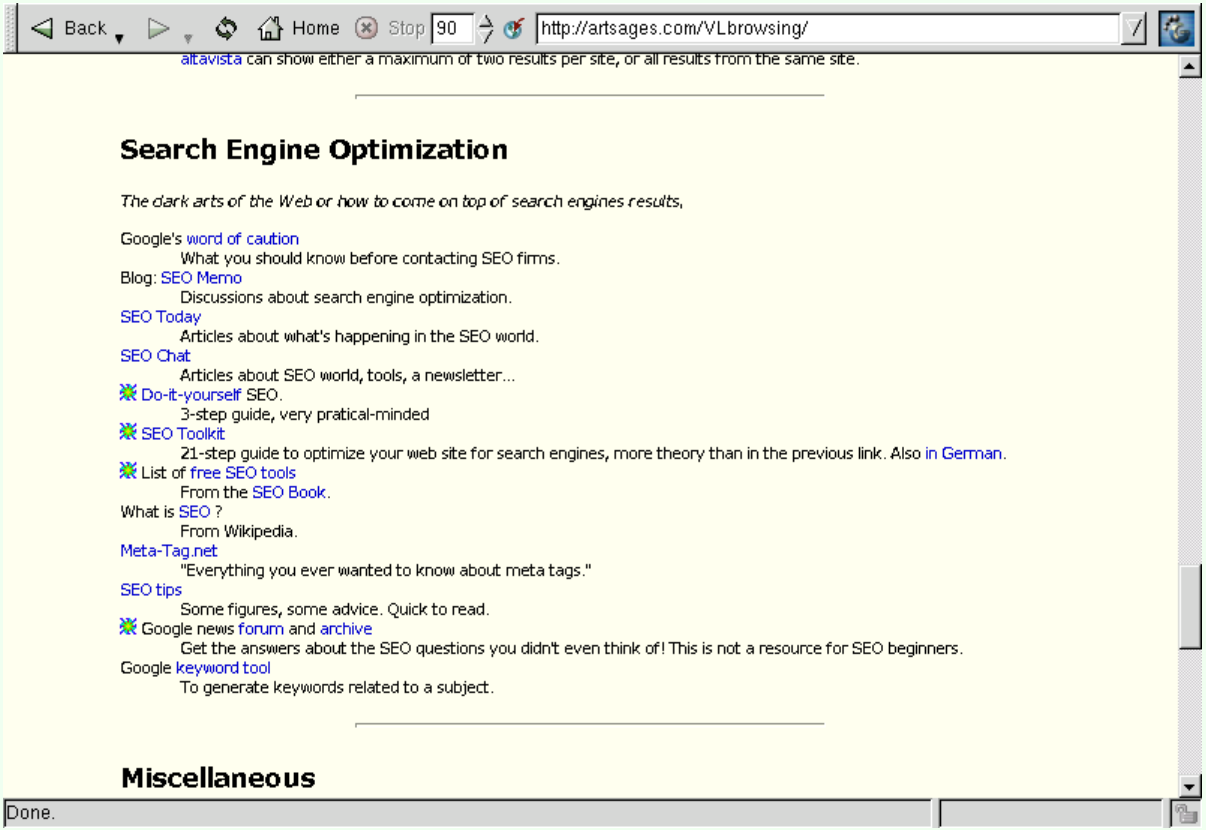
Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

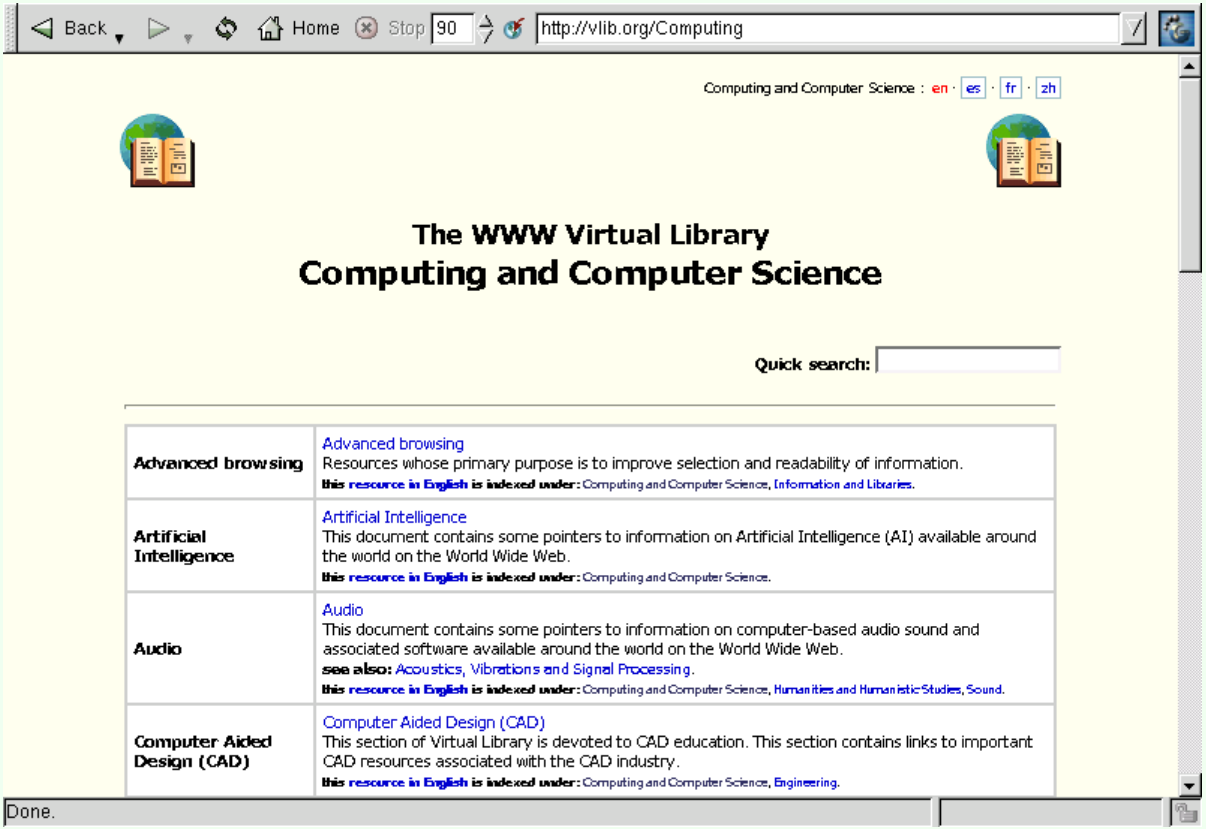


Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Equally one could have reached this from *computing*:



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

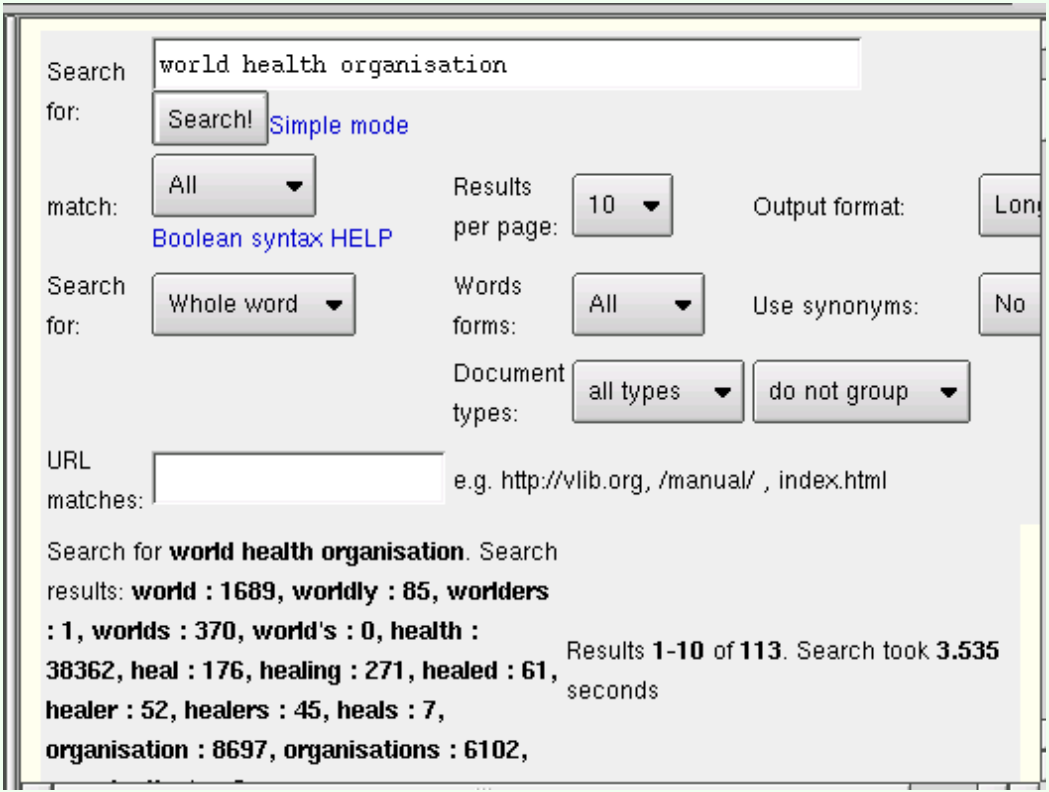
World Health Organisation —simple search

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

World Health Organisation —“string” search

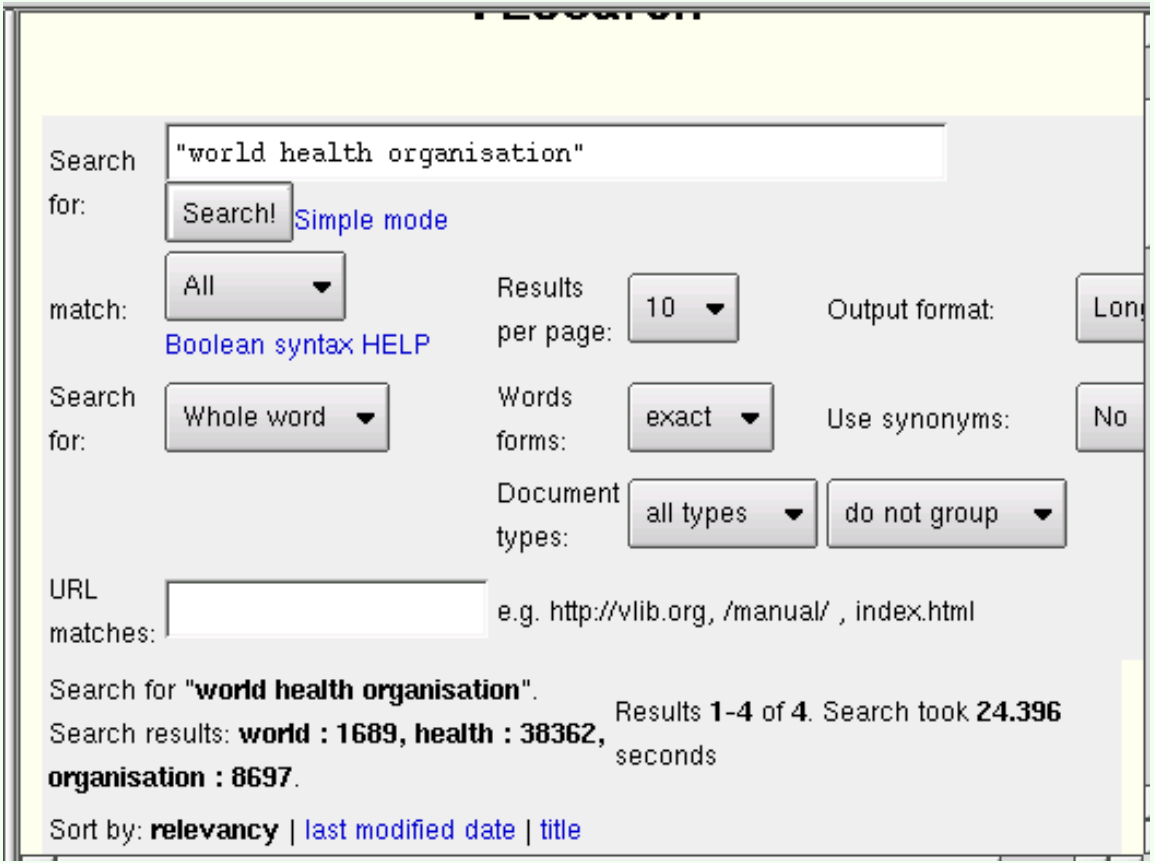
Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

World Health Organisation —search excluding synonyms (organisation/organization)



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

World Health Organisation —search for the “string” without synonyms



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

World Health Organisation:

simple search	179
“string”	10
without synonyms	113
“string” without synonyms	4

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Usage:

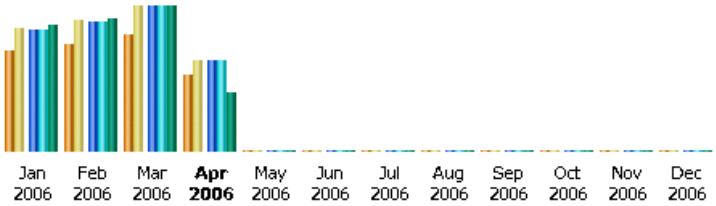
Summary

Reported period	Month Apr 2006
First visit	01 Apr 2006 - 00:02
Last visit	24 Apr 2006 - 02:01

	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Viewed traffic *	**16350**	**19539** (1.19 visits/visitor)	**53071** (2.71 pages/visit)	**53071** (2.71 hits/visit)	**191.51 MB** (10.03 KB/visit)
Not viewed traffic *			**7680**	**7680**	**10.14 MB**

* Not viewed traffic includes traffic generated by robots, worms, or replies with special HTTP status codes.

Monthly history



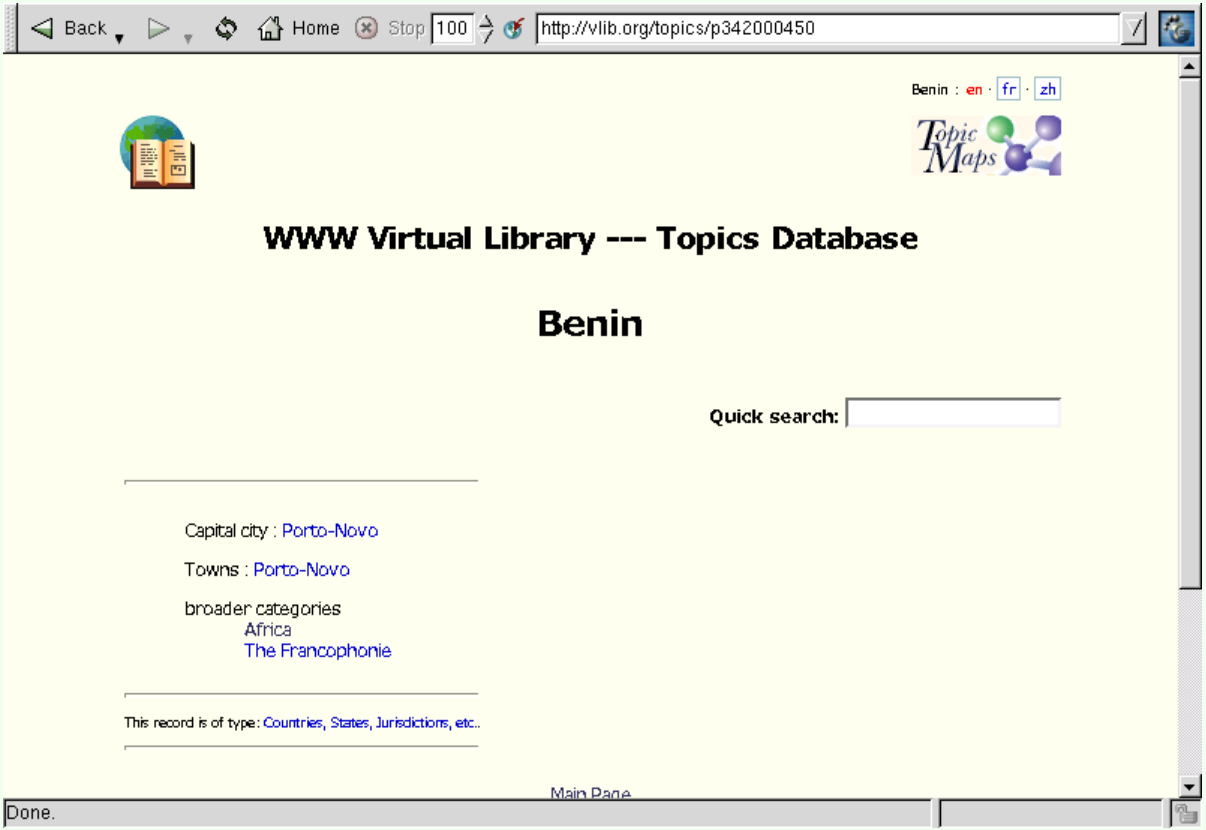
Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Search Engine of member VL pages, ‘VLsearch’:

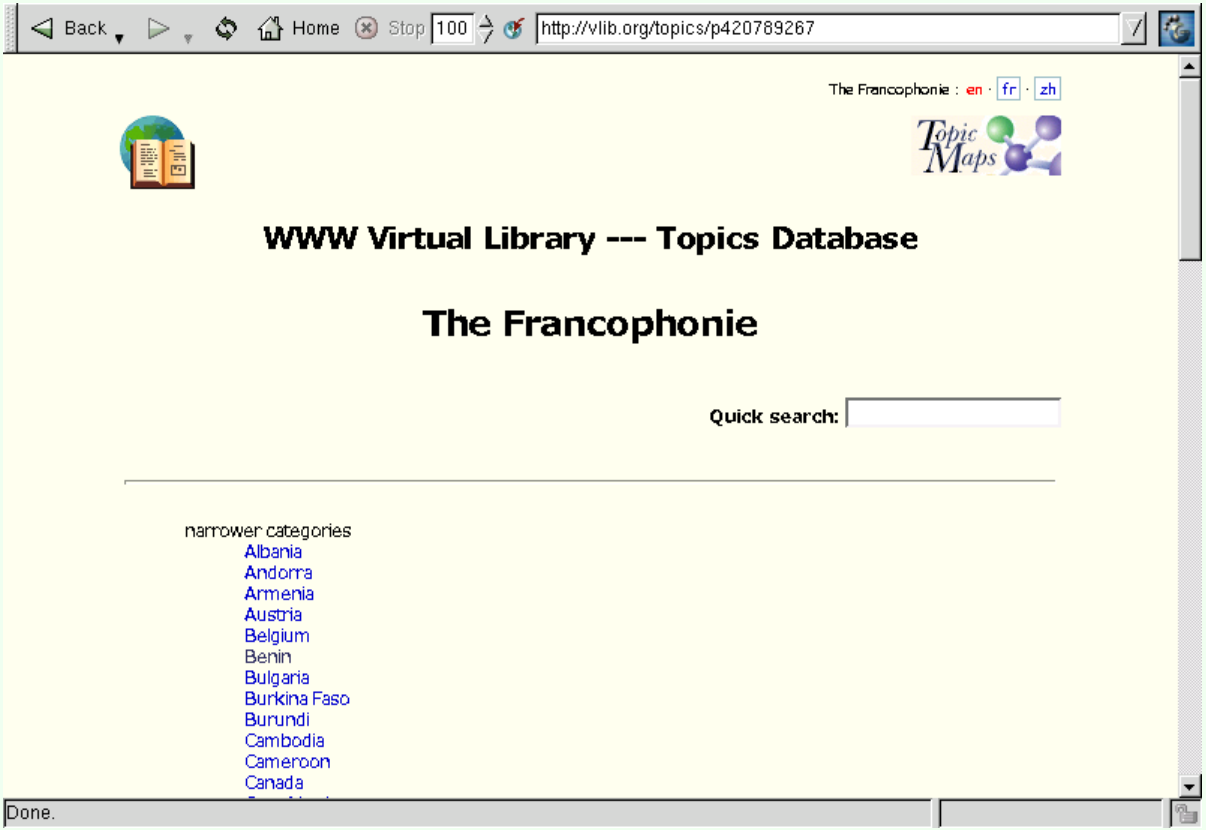
- relatively young (just over one year old)
- still adding member VLs and ensuring their coverage
- popular: twenty to thirty thousand visitors/month
- some 10% of users are using a non-English language interface

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

3.2.3. Topic Maps



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion



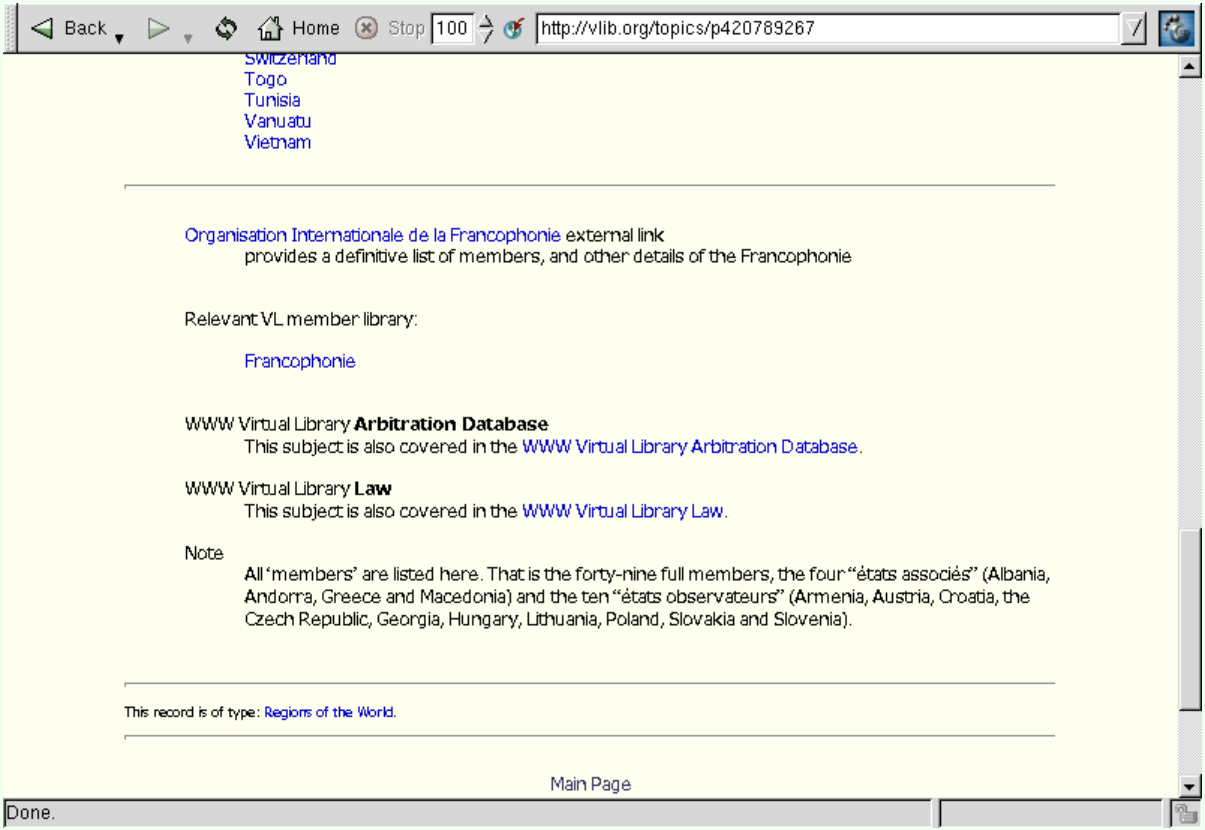
Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion



Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

A Topic Map links TOPICS by ASSOCIATIONS.

A thesaurus is a simple example. It has very few association types (e.g. BROADER TERM and NARROWER TERM).

In a TM associations can be anything describable. They are also inherently, themselves topics.

Thus associations such as ‘is the author of’, ‘wrote the work’, ‘is published by’, ‘is the publisher of’.

The ability to query the *associations* between topics as well as the (page) topics themselves is a very powerful tool.

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

Topic mapping topics in member VLs:

Topic Maps originated from research on the meta-indexing of textbooks –they fit very closely with the needs of the WWWVL.

- relatively young (one year old)
- experimental
- fairly obviously, works best with easily defined/fixed topics (countries, languages, . . .)
- Much to do:
  - The creation of a TMQL (TOPIC MAP QUERY LANGUAGE) would release the power of the knowledge contained in topic maps –at the moment our present pages probably appear to most users as just an over-sized hierarchy with rather good cross-linking
  - if we are to achieve a “semantic Web” this would be a major contribution towards it

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

3.3. Infrastructure

Linux

The VL central services were put on an existing, running server in Geneva in late 2004.

Although orders of magnitude different from the server's previous usage, as one might expect there were no significant problems.

Webserver: *Apache*
Database (TM) software: 'bespoke' (but quite simple Perl).

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

4. The WWWVL and the Future

Volatility

both ‘loss’ and having ‘had the consultants in’ . . .

What to do about loss:

to cache / to archive ?

greater linking to archive.org ?

. . . ?

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

Censorship/Partisanship

ordre public

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Censorship/Partisanship

ordre public

sites/domains that advocate/condone child prostitution

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

Censorship/Partisanship

ordre public

sites/domains that advocate/condone child prostitution
sites/domains that advocate/condone torture . . . ?

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

Censorship/Partisanship

ordre public

sites/domains that advocate/condone child prostitution
sites/domains that advocate/condone torture . . . ?

Practically a VL on, say, Classical Music should not be advocating or encouraging classical music; it should be giving a dispassionate, balanced review of the best web resources available on its subject area.

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

The challenges the Virtual Library faces in relation to knowledge management, in the cataloging of resources and making an accessible, effective interface to those resources are far from unique.

They are major challenges that will face us all, if in various forms, over the coming decades.

Historical Introduction

Ethos

The WWWVL in 2006

The WWWVL and the . . .

Conclusion

5. Conclusion

- Creation of a stable management structure, building on a decade-and-a-half of remarkable voluntary administration.
- Whilst guarding the primary role of knowledge dissemination, widening the target population.

  The VL has always prided itself on using small (fast loading) pages, 'frames' free and with only essential graphics content. What was (and is) of advantage to those with poor connectivity, is now relevant for the increasing number of small screen devices.
- The central catalogue is now available in English and French, Spanish is in the process of being loaded. An experimental Chinese interface is already in operation and is being re-edited.
  The search engine has Arabic, English, French and Spanish interfaces. It is intended to have both servers working in all five of these languages as well as Russian.
- 'Outreach' both to members (a first regional meeting in 2006) and to non-members (*Linux Days*, the *Topic Maps Research and Applications* conference, the *4^{me} Rassem-*

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the...
Conclusion

blement pour les droits de l'homme, today) has broadened interest in the VL as well as providing valuable feedback from experts in related fields.
Such activities will need expanding: to attract new members (and thus improve coverage); to inform users and potential users; to attract sponsors and supporters; and, not least, to learn from all these persons.

- Knowledge can only be effectively disseminated if it can be found. A hierarchical catalogue and a VL specific search engine are both essential parts of this. Mapping subjects and their semantic linkages is an area the VL is developing using topic maps and published subject/resource indicators (PSIs).
- Such a mapping should link deeper into member virtual libraries.
- Amongst central services to members is the checking of the validity of links on their pages. Extending this to a robust quality audit including link checking, page accessibility criteria (i.e. WAI) and spellchecking is an important

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the...
Conclusion

goal.
A detailed booklet of guidelines has already been published and distributed to members.

- Feedback, particularly today, feedback on how to improve knowledge dissemination in parallel with ‘development’ initiatives, is always sought, and valued.

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion

vlib.org

Historical Introduction
Ethos
The WWWVL in 2006
The WWWVL and the . . .
Conclusion