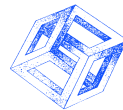


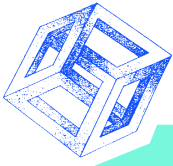
# Internet

## Recherches de données sur internet

William El Kaim



<http://www-masi.ibp.fr/~William.El-Kaim>



# Trouver l'information sur Internet

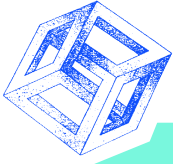
- + La vraie valeur ajoutée de l'internet c'est l'accès et la recherche de l'information 'gratuitement'.**
- + Unité de temps, de lieu et de contenu.**
- + Univers virtuel spécialisé sur un sujet**

---

## OUTILS ET PROTOCOLES

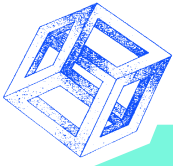
---

- + Archie : recherche de logiciels.**
- + Wais : Recherche dans des bases de données (ancien)**
- + Glimpse : recherche dans des bases de données (récent)**
- + Search Engine : retrouver des documents sur le WWW**



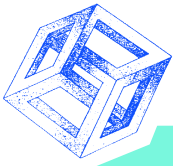
# Les problèmes

- + Recherche par la découverte.**
- + Les documents sont sous différentes formes et dans différentes langues.**
- + Le temps d'accès à l'information est différent.**
- + La qualité des documents ne peut-être évaluées qu'après réception.**
- + Les protocoles de recherche et les outils adaptés sont très nombreux.**
- + L'accès est navigationnel => on peut "surfer" et être attiré par de nouveaux liens vers des sujets en dehors du cadre de la recherche.**

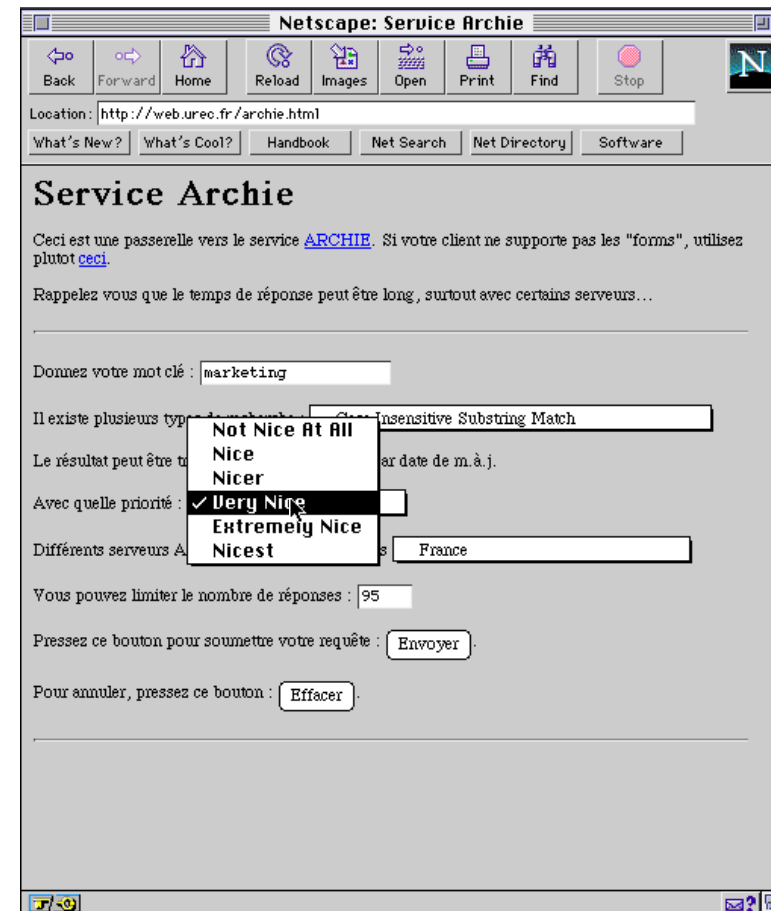


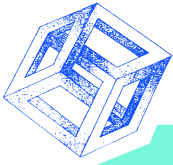
# Archie : Recherche de fichiers

- + Recherche de fichiers et pas des documents par mots clés.
- + Archie peut-être interrogé par Telnet, E-mail et WWW.
- + Il existe des serveurs Archie en France (Rennes I).
- + Le tri des réponses par date, par serveur ou par pays.
- + le mode de recherche : exact, avec le mot contenu dans un autre, ou une expression régulière.
- + Les expressions régulières : commence par (^), finit par (\$), plusieurs occurrences (\*) et prise en compte de certains caractères [ ] ). `^chess[0-9]` => `chess206.zip` et `chess207.zip` sont possibles.

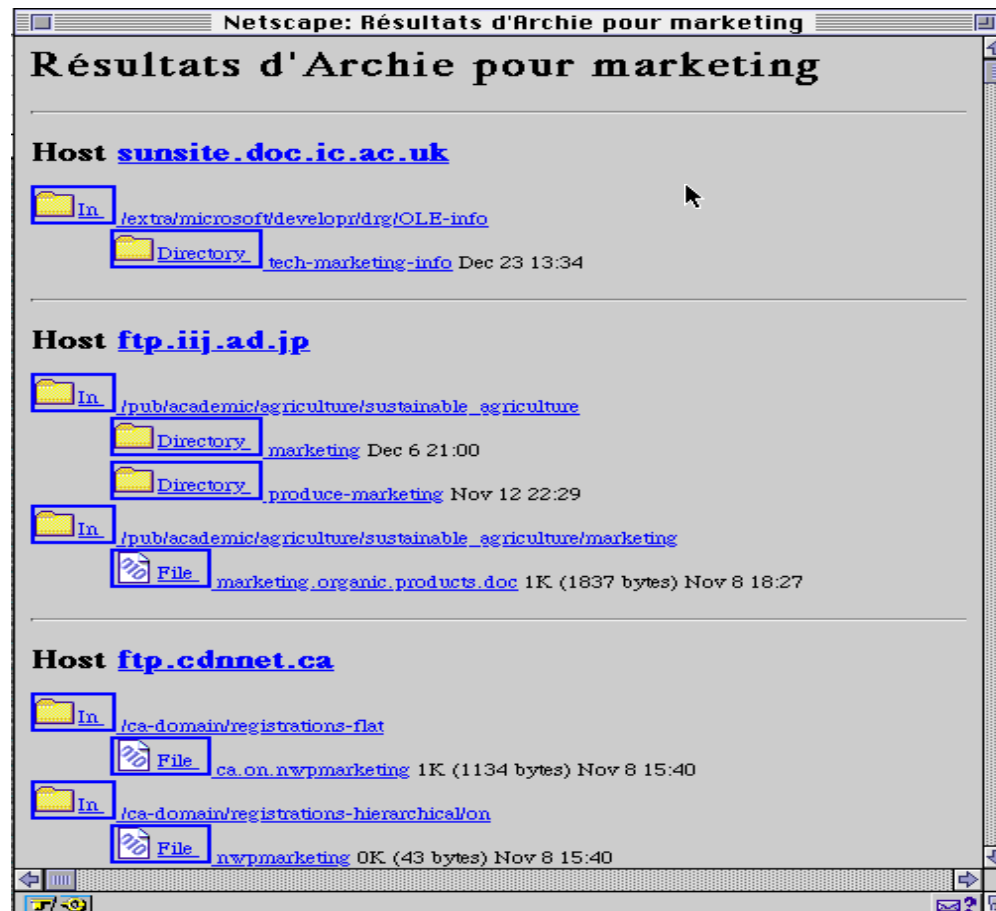


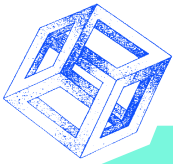
# Archie : exemple





# Archie (résultat)

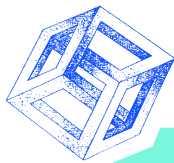




# WAIS



- + **Wide Area Information Servers fut la première base de données textuelle sur Internet.**
- + **Wais est basée sur une indexation des documents à rechercher. La recherche se fait par mot clef et est classée en fonction du degré de pertinence.**
- + **Il est possible de visualiser et de récupérer les documents trouvés.**




# DEA - DESS au ministère (Wais)

**Netscape: LISTE DE DOMAINES**

Location:

What's New | What's Cool? | Handbook | Net Search | Net Directory | Newsgroups

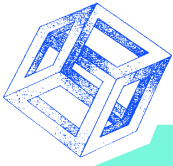
## Liste de domaines



Vous recherchez un DEA.  
Pour le trouver, vous pouvez:

- Choisir l'un des domaines suivants:
  - [MATHEMATIQUES ET LEURS APPLICATIONS](#)
  - [SCIENCES DE LA MATIERE ET GENIE DES PROCEDES](#)
  - [SCIENCES DE LA TERRE ET DE L'UNIVERS, ESPACE](#)
  - [SCIENCES ET TECHNOLOGIE DE L'INFORMATION](#)
  - [BIOLOGIE, MEDECINE, SANTE](#)
  - [SCIENCES HUMAINES ET SOCIALES](#)
  - [SCIENCES JURIDIQUES, POLITIQUES ECONOMIQUES ET DE GESTION](#)
  - [MECANIQUE, GENIE ELECTRIQUE, PRODUCTION, TRANSPORTS ET GENIE CIVIL](#)
- Choisir une région sur la [CARTE DE FRANCE](#).
- Indiquer un ou plusieurs mots ou débuts de mot qui décrivent le sujet ou le lieu du DEA:





# Résultat (Wais)

The screenshot shows a Netscape browser window titled "Netscape: LISTE DE DEA". The address bar contains the URL "http://dgrt.mesr.fr/SIF3-bin/recherche.pl/0/0/-SIF3-dspt.html". Below the address bar are navigation buttons: Back, Forward, Home, Reload, Images, Open, Print, Find, and Stop. There are also buttons for "What's New?", "What's Cool?", "Handbook", "Net Search", "Net Directory", and "Newsgroups".

The main content area displays the following text:

Dans la région:  
**France**  
dans le domaine:  
**TOUTES DISCIPLINES**

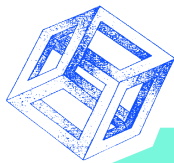
Vous avez choisi les DEA dont le titre, la discipline, l'établissement ou la ville comporte le ou les mots suivants:  
**INFORMATIQUE**

There are three small icons: a green arrow pointing left, a question mark, and the flag of the United Kingdom.

Vous pouvez:

- changer de mots
- choisir un DEA dans la liste qui suit:

- [INFORMATIQUE](#)  
Informatique  
UNIVERSITE AIX-MARSEILLE 2
- [PRODUCTIQUE ET INFORMATIQUE](#)  
Automatique, robotique  
UNIVERSITE AIX-MARSEILLE 3
- [INFORMATIQUE FONDAMENTALE](#)  
Informatique  
UNIVERSITE AMIENS



## Suite ... (Wais)

DEA:

### INFORMATIQUE

Discipline principale:

#### Informatique

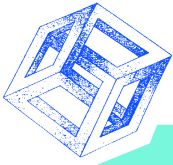
(SCIENCES ET TECHNOLOGIE DE L'INFORMATION)



Responsable: Mr Michel VAN CANEGHEM  
Ville Universitaire: MARSEILLE

Etablissement principal: [UNIVERSITE AIX-MARSEILLE 2](#)  
Adresse: Jardin du Pharo  
: 58 boulevard Charles-Livon  
: 13284 MARSEILLE CEDEX 07

Adresse du DEA: UNIVERSITE AIX-MARSEILLE 2  
: 163 avenue de Luminy  
:  
: 13288 MARSEILLE CEDEX  
Tél: 91-26-93-16  
Fax: 91-26-92-75



# Glimpse

**Computer Science Bibliography Glimpse Server**

This experimental server gives you flexible search capabilities using *Glimpse*.

Click here - [Glimpse Working Group - University of Arizona, Computer Science Dept](#) - for more information about Glimpse, which is available for personal UNIX file systems as well as HTTP servers. Some information is also given below as well as instructions and tips on how to form search queries.

The underlying database is a [large collection of CS bibliographies](#), all in bibTeX format, put together by Alf-Christian Achilles.

A search here will cover the whole database (about 300,000 citations). You can follow any of the links below to search only the corresponding database. If your browser doesn't support forms, you can still [do the searching](#).

What would you like to search for?  AND  AND

Case insensitive  Allow partial match to a word

Number of misspelings allowed:

Maximum number of bibliography files returned:

Maximum number of matched bibliographies per file returned:

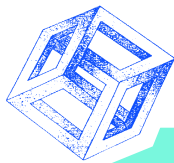
Press this button to submit the query:

To reset the form, press this button:

---

**Areas**

- [Abbreviations](#)
- [Artificial Intelligence](#)
- [Compiler Technology and Type Theory](#)
- [Databases](#)
- [Distributed Systems and Networking](#)
- [Computer Graphics and Vision](#)
- [\(Computational\) Mathematics](#)
- [Neural Networks](#)



# Glimpse (2)

**Netscape: Computer Science Bibliography Glimpse Server**

## Computer Science Bibliography Glimpse Server

This experimental server gives you flexible search capabilities using *Glimpse*.

Click here - [Glimpse Working Group - University of Arizona, Computer Science Dept](#) - for more information about Glimpse, which is available for personal UNIX file systems as well as HTTP servers. Some information is also given below as well as instructions and tips on how to form search queries.

The underlying database is a [large collection of CS bibliographies](#), all in bibTeX format, put together by Alf-Christian Achilles.

A search here will cover the whole database (about 300,000 citations). You can follow any of the links below to search only the corresponding database. If your browser doesn't support forms, you can still [do the searching](#).

What would you like to search for?  AND  AND

Case insensitive  Allow partial match to a word

Number of misspellings allowed:

Maximum number of bibliography files returned:

Maximum number of matched bibliographies per file returned:

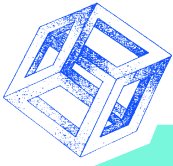
Press this button to submit the query:

To reset the form, press this button:

---

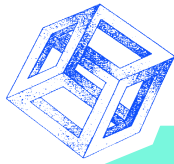
### Areas

- [Abbreviations](#)
- [Artificial Intelligence](#)
- [Compiler Technology and Type Theory](#)
- [Databases](#)
- [Distributed Systems and Networking](#)
- [Computer Graphics and Vision](#)
- [\(Computational\) Mathematics](#)
- [Neural Networks](#)

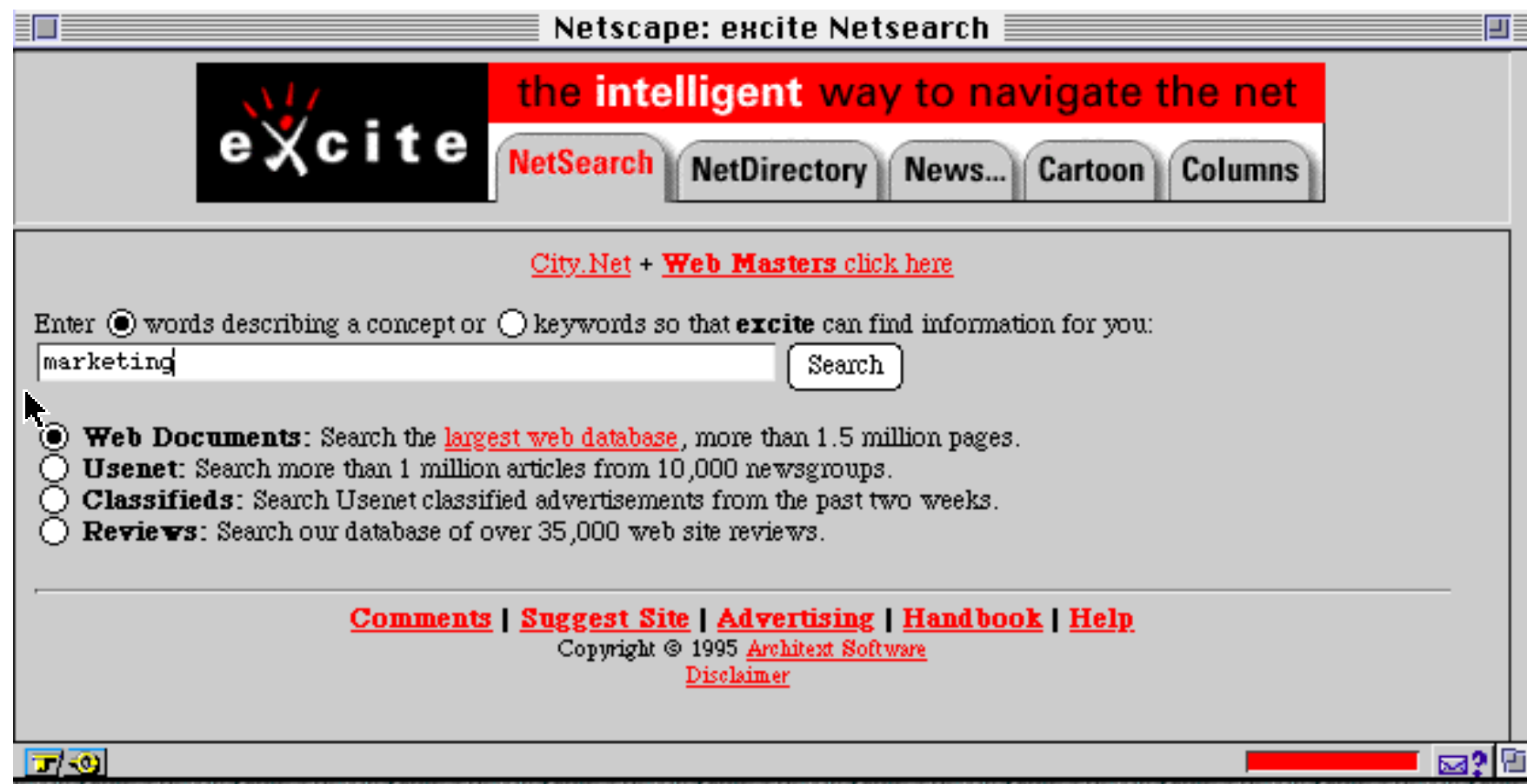


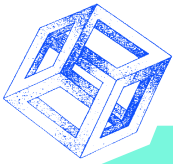
# Les moteurs de recherche sur WWW

- + **Ecila**
- + **Excite**
- + **Fulcrum**
- + **Galaxy Einet**
- + **GNN**
- + **Harvest**
- + **Infoseek**
- + **Inktomi**
- + **Lycos**
- + **Magellan (Mc Kinley)**
- + **Netmall**
- + **Nikos**
- + **Opentext**
- + **Starting Point**
- + **Veronica**
- + **Wandex**
- + **Web Crawler**
- + **What's new too**
- + **World-Wide Web Servers**
- + **World Wide Web Worm**
- + **W3 Search Engines**
- + **Yahoo**



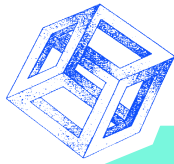
# Mon préféré : Excite





# Excite (2)





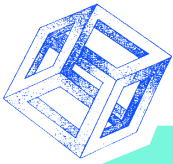
# Excite (3)

**TRADE-Net Other Hot Web Links...**

- [Harvard Business School - Marketing Area](#)
- [Internet Marketing Inc.](#) - offer deals for businesses who want their presence felt on the Information Superhighway.
- [On-line Marketing](#)
- [Ad Methods Halpern and Eurich Advertising, Inc.](#) - A full service advertising and marketing firm with over 20 years experience.
- [adfx - Virtual Advertising](#) - adfx - Virtual Advertising provides low cost access to commercial information display on the World Wide Web. A contemporary advertising agency pursuing a global community while being a doorway for businesses, entrepreneurs and individuals deciding to present a message to the Internet. Devoted to Electronic Media Marketing in all it's forms.
- [Admarket](#) - gateway to marketing, media, advertising and public relations content
- [AfterHours Communications Corp.](#) - AfterHours is a high-tech advertising and marketing agency that specializes in the small startup.
- [Black Box](#) - Design Marketing Advertising
- [CouponNet](#) - provides an international internet presence for manufacturers and retailers who distribute coupons and rebates to consumers as part of their interactive advertising and marketing strategy.
- [Floathe Johnson High-Tech Advertising](#) - Find out about effective marketing for technology companies. You can get details on the advertising, public relations, new media and TechnoBranding services that Floathe Johnson offers to help make its clients more successful.
- [Hensley Segal Renschler](#) - A marketing communications firm.
- [IMPULSE!](#) - is a Electronic Marketing and Advertising Production Company. Specializes in creating exciting program length advertisements
- [Liggett Stashower](#) - a full service marketing communications company.
- [McMonigle & Spooner](#) - an advertising, marketing and design firm providing a fresh approach to traditional and non-traditional marketing and design challenges.
- [Modem Media](#) - an advertising and marketing agency dedicated exclusively to interactive media. We have executed programs on every interactive platform ranging from interactive telephone voice response, to online services, the Net, CD-ROM, multimedia production, interactive fax, videogames, interactive TV and virtual reality.
- [Origin Communications, Inc.](#) - Origin Communications is an advertising firm specializing in high-tech, high-end marketing communications.
- [Richardson, Myers & Donofrio, Inc.](#) - information on social marketing, tour and travel, business to business, as well as case studies.
- [Tom Davis+Company](#) - Provides advertising, design, and marketing services with both business-to-business and direct-to-consumer objectives.
- [Wahlstrom & Company](#) - Directory Marketing Specialists -- Leading advertising agency specializing in National Yellow Pages.
- [CKS Partners Ltd.](#) - Integrated Marketing Communications
- [wolfBayne Communications](#) - emphasizes strategic and tactical public relations, marketing communications management and logistics, and Internet marketing development and implementation.
- [Web Digest for Marketers](#) [new] - bi-weekly bulleting reporting on the latest marketing sites on the web.
- [World Wide Web Aerospace Business Development Center](#) - A starting point for aerospace marketing, procurement and business development professionals navigating the World Wide We
- [Right Angle Design](#) [new] - professional web publishing, marketing and graphic design for architects and developers.
- [Museum Web](#) - a company specializing in the Marketing & Communications of Museums, the Arts, Antiques, and Valuables & Collectibles fields. We create WWW Home Pages, Sites and Servers for the above types of Institutions and Companies. We also

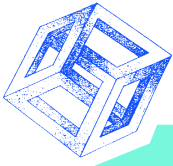
http://coupon.com/coupon.html





# Comment faire connaître son site





# Submit-it (2)

The screenshot shows a Netscape browser window with the title "Netscape: Submit It!". The page contains a form for submitting URLs to various search engines. At the top, there is an "Email address:" field and a "Register" button. Below this is a grid of 12 small icons representing different search engines. Underneath the grid is a list of search engines with checkboxes next to their names. The form includes fields for "Title of your site:", "URL:" (with "http://" pre-filled), "Category:" (with "Health & Fitness" selected), "Keywords:" (separated by spaces), and "Name of Organization or Business:" (Use your name if personal site). The browser's status bar at the bottom shows the URL "http://webcrawler.com/WebCrawler/SubmitURLS.html".

Email address:  Register

Yahoo  WebCrawler  Starting Point  Lycos  
 Infoseek  What's New Too!  Galaxy  Harvest  
 Whole Internet Catalog  Apollo  
 LinkStar  New Rider's WWW YP  Nerd World Media

**Title of your site:**

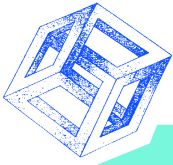
**URL:**

**Category:**  
 If **Other**, please Specify

**Keywords:** (separated by spaces)

**Name of Organization or Business:** (Use your name if personal site)

http://webcrawler.com/WebCrawler/SubmitURLS.html



# Submit-it (3)

**Netscape: Submit It!**

**URL:**  
http://

**Category:**  
Health & Fitness If **Other**, please Specify

**Keywords:** (separated by spaces)

**Name of Organization or Business:** (Use your name if personal site)

**City:**

**State/Province:** (must use two-letter code if in US)

**Country:**  
US

**Contact Name:** (Your Name)

**Contact E-Mail Address:** (Your E-mail Address)

**Description:** (no HTML tags please)  
A single paragraph (25 words or less) that briefly describes your site.

OK, move on to the submitting area No, reset this form

6% of 14K (at 11 bytes/sec)