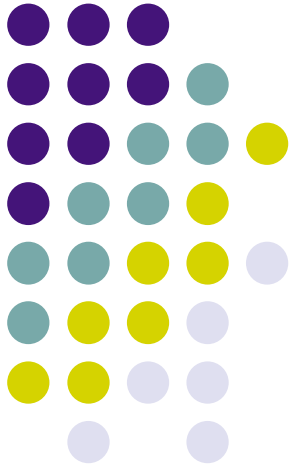
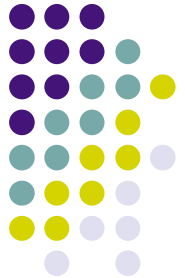
Research using the Internet

Computer Literacy 1 lecture 5

30/09/08

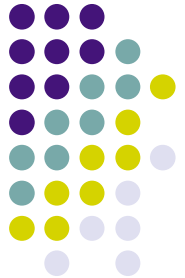


Topics



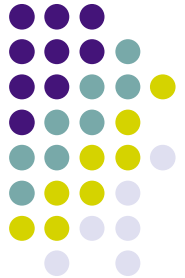
- Topics
  - Internet
  - The Web
  - Search engines
  - Classified directories
  - Evaluating data

Aims



- Know which search tool to use
- Able to access quality of sources

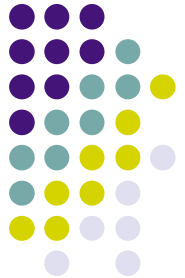
The Internet



- The internet is a system of interconnected computer networks.
  - Formed by thousands of networks run by businesses, governments, universities etc.
  - Internet allows different users to communicate e.g. using email or instant messaging (chatting)
  - Another possibility to communicate is using the World Wide Web (www) by creating own website.

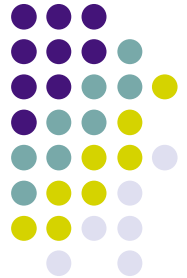
The Internet

The Web



- The WWW is a distributed browsing and searching system originally developed at CERN by Tim Berners-Lee
- The web was started in 1990 as a way for users to share access to files
- The files that can be accessed are known as documents
- The programs that allow you to display the files on your computer are known as browsers.

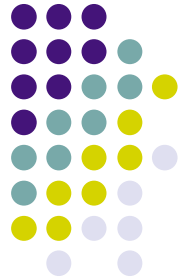
Hypertext



- Webpages are written in format that allows to be read by browsers
  - The contend is annotated with information about how the contend should be displayed
- One key in feature of the annotation is that a word in one page can be *linked* to another webpage
- Words annotated in this way are called **HYPERTEXT**

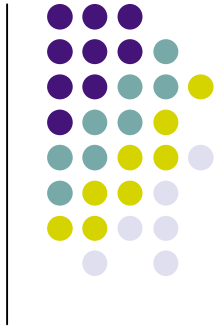
http://en.wikipedia.org/wiki/Hypertext

HTML and HTTP



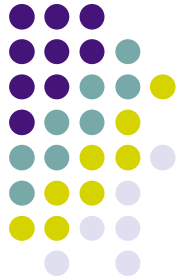
- The language used for this annotation is **H**yptertext **M**arkup **L**anguage ➔ HTML
- Documents linked together using hypertext are called websites
- Computers that store websites are called Web Servers
- Each website has a unique address code called a URL (uniform resource locator) ➔ browsers need this information to access webservers

HTML and HTTP



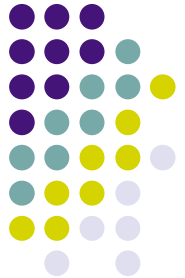
- HTTP = **H**ypertext **T**ransfer **P**rotocol
  - Is a web standard protocol for how browsers and servers communicate
    - Example: The URL for the University of Edinburg is http://www. ed.ac.uk, where http is the transfer and www.ed.ac.uk is the domain specific address of the host containing the information

Features of the Web



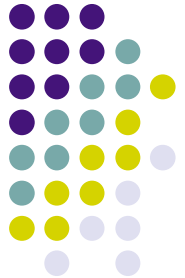
- Huge variety of information sources
- Available 24 hours
- Can be up-to-date
- Accessible from any computer

The Web



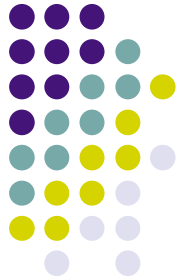
- It's unregulated
  - Free from censorship (matter of opinion)
  - Little quality control
  - No standard vocabulary
- Sometimes difficult to search

Search Engines



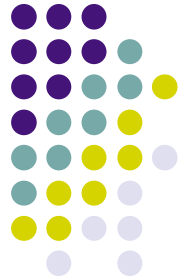
- Build databases of indexed websites
  - Automatically created
  - Normally don’t cover the whole web
  - Can’t access databases
  - No quality control
- Search by keywords only
  - Results ranked by relevance and your preferences

Search engines



- General queries
  - Generate many hits
  - Often with low relevance
- Good for
  - Specific queries e.g. particular people or organisations
  - Finding lots of information

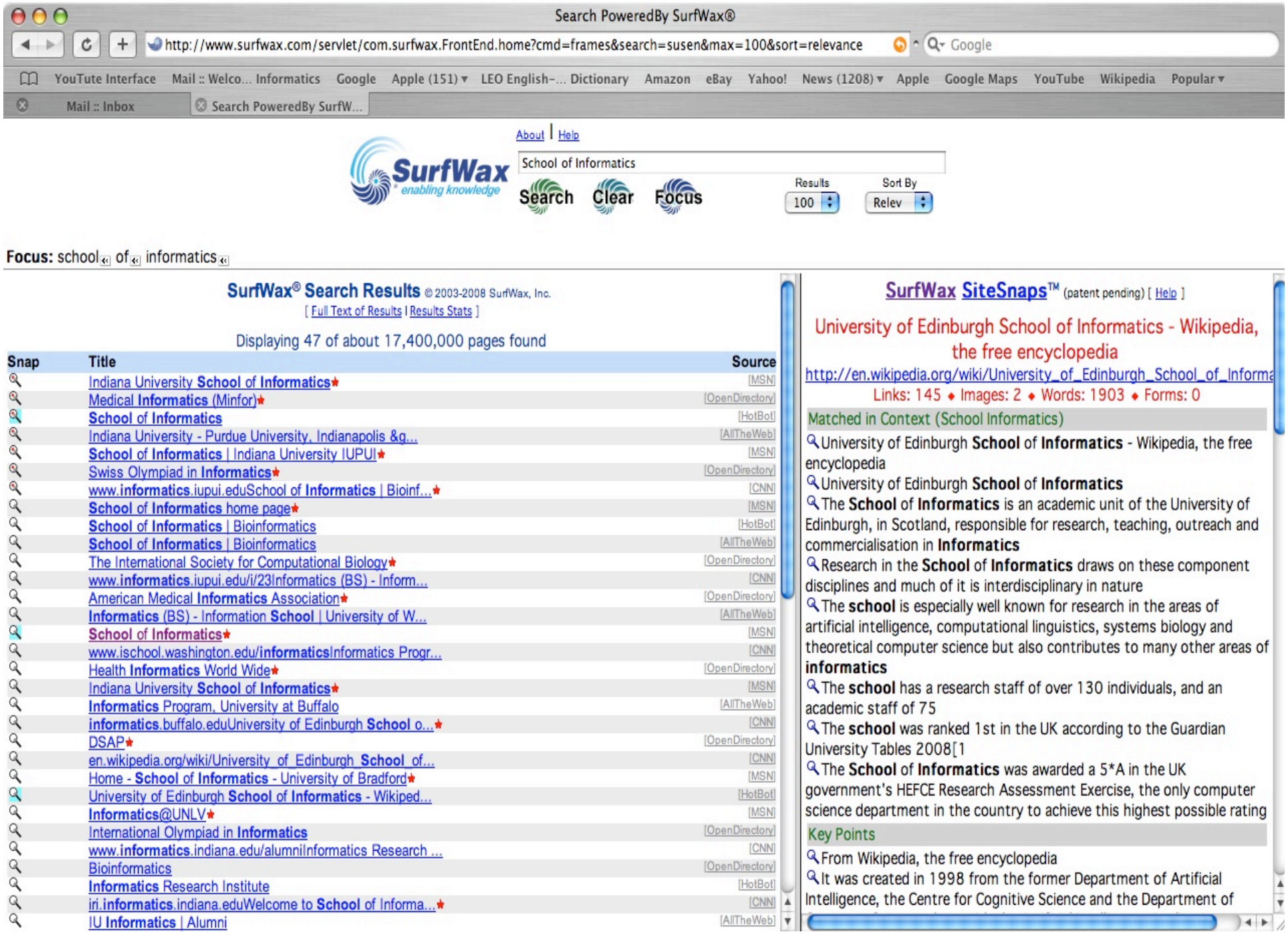
Examples of Metasearch Engines



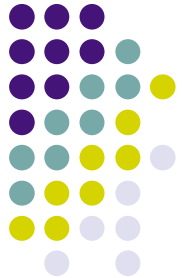
- Surfwax has 2 interesting interface features
  - Information about the author
  - Visual interface to help assess the relevance of the multitude of hits

  http:www.surfwax.com

Other metasearch engines include
Metacrawler, DogPile, Webcrawler, HotBot


Search PoweredBy SurfWax®
http://www.surfwax.com/servlet/com.surfwax.FrontEnd.home?cmd=frames&search=susen&max=100&sort=relevance
Google
YouTute Interface
Mail :: Welco... Informatics
Google
Apple (151)
LEO English-... Dictionary
Amazon
eBay
Yahoo!
News (1208)
Apple
Google Maps
YouTube
Wikipedia
Popular
Mail :: Inbox
Search PoweredBy SurfW...
About | Help
SurfWax
enabling knowledge
School of Informatics
Search
Clear
Focus
Results
100
Sort By
Relev
Focus: school of informatics
SurfWax® Search Results © 2003-2008 SurfWax, Inc.
[Full Text of Results | Results Stats]
Displaying 47 of about 17,400,000 pages found
Snap
Title
Source
Indiana University School of Informatics [MSN]
Medical Informatics (Minfor) [OpenDirectory]
School of Informatics [HotBot]
Indiana University - Purdue University, Indianapolis &g... [AllTheWeb]
School of Informatics | Indiana University IUPUI [MSN]
Swiss Olympiad in Informatics [OpenDirectory]
www.informatics.iupui.eduSchool of Informatics | Bioinf... [CNN]
School of Informatics home page [MSN]
School of Informatics | Bioinformatics [HotBot]
School of Informatics | Bioinformatics [AllTheWeb]
The International Society for Computational Biology [OpenDirectory]
www.informatics.iupui.edu/i/23Informatics (BS) - Inform... [CNN]
American Medical Informatics Association [OpenDirectory]
Informatics (BS) - Information School | University of W... [AllTheWeb]
School of Informatics [MSN]
www.ischool.washington.edu/informaticsInformatics Progr... [CNN]
Health Informatics World Wide [OpenDirectory]
Indiana University School of Informatics [MSN]
Informatics Program, University at Buffalo [AllTheWeb]
informatics.buffalo.eduUniversity of Edinburgh School o... [CNN]
DSAP [OpenDirectory]
en.wikipedia.org/wiki/University_of_Edinburgh_School_of... [CNN]
Home - School of Informatics - University of Bradford [MSN]
University of Edinburgh School of Informatics - Wikiped... [HotBot]
Informatics@UNLV [MSN]
International Olympiad in Informatics [OpenDirectory]
www.informatics.indiana.edu/alumniInformatics Research ... [CNN]
Bioinformatics [OpenDirectory]
Informatics Research Institute [HotBot]
iri.informatics.indiana.eduWelcome to School of Informa... [CNN]
IU Informatics | Alumni [AllTheWeb]
SurfWax SiteSnaps™ (patent pending) [Help]
University of Edinburgh School of Informatics - Wikipedia, the free encyclopedia
http://en.wikipedia.org/wiki/University_of_Edinburgh_School_of_Informa
Links: 145 ♦ Images: 2 ♦ Words: 1903 ♦ Forms: 0
Matched in Context (School Informatics)
University of Edinburgh School of Informatics - Wikipedia, the free encyclopedia
University of Edinburgh School of Informatics
The School of Informatics is an academic unit of the University of Edinburgh, in Scotland, responsible for research, teaching, outreach and commercialisation in Informatics
Research in the School of Informatics draws on these component disciplines and much of it is interdisciplinary in nature
The school is especially well known for research in the areas of artificial intelligence, computational linguistics, systems biology and theoretical computer science but also contributes to many other areas of informatics
The school has a research staff of over 130 individuals, and an academic staff of 75
The school was ranked 1st in the UK according to the Guardian University Tables 2008[1
The School of Informatics was awarded a 5*A in the UK government's HEFCE Research Assessment Exercise, the only computer science department in the country to achieve this highest possible rating
Key Points
From Wikipedia, the free encyclopedia
It was created in 1998 from the former Department of Artificial Intelligence, the Centre for Cognitive Science and the Department of

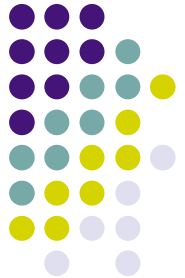
Classified Directory



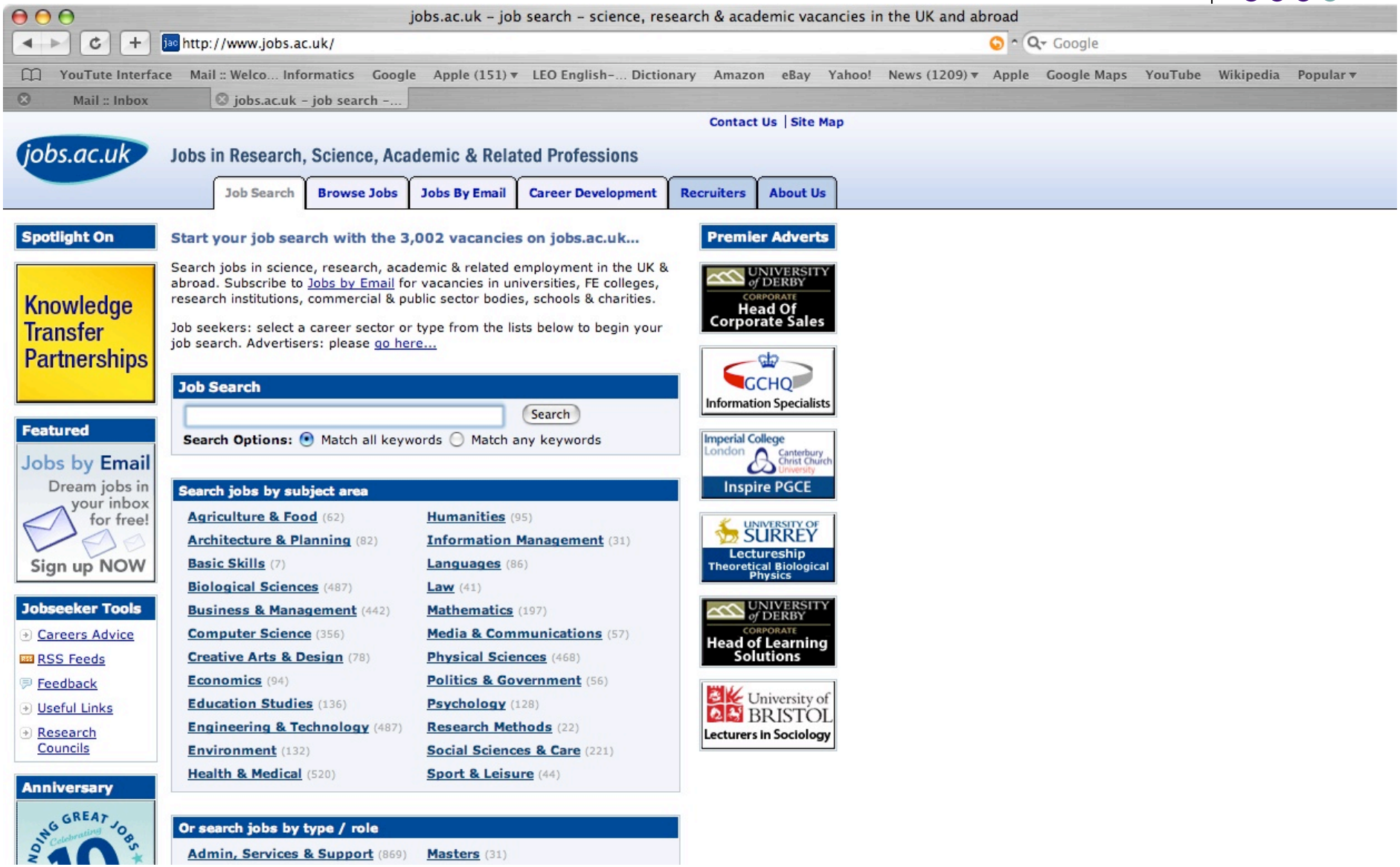
- Database of websites collected manually
  - Records organised systematically
  - Some quality check
  - Only records of websites are searched
  - Smaller web coverage then engines
- Good for general enquiries
  - Allows browsing
  - Wide coverage of subjects

  Example: http://dir.yahoo.com/

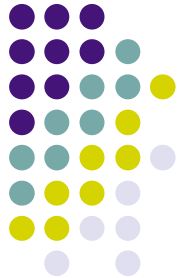
Subject gateways



- Classified directories for subject areas
  - For higher education (you!)
  - High quality
  - Good coverage of subject
  - Links
- Good for browsing a subject area
  - Get reliable information

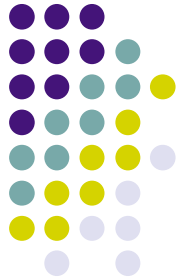

jobs.ac.uk – job search – science, research & academic vacancies in the UK and abroad
http://www.jobs.ac.uk/
Google
YouTute Interface
Mail :: Welco... Informatics
Google
Apple (151)
LEO English-... Dictionary
Amazon
eBay
Yahoo!
News (1209)
Apple
Google Maps
YouTube
Wikipedia
Popular
Mail :: Inbox
jobs.ac.uk – job search -...
Contact Us | Site Map
jobs.ac.uk
Jobs in Research, Science, Academic & Related Professions
Job Search
Browse Jobs
Jobs By Email
Career Development
Recruiters
About Us
Spotlight On
Knowledge Transfer Partnerships
Featured
Jobs by Email
Dream jobs in your inbox for free!
Sign up NOW
Jobseeker Tools
Careers Advice
RSS Feeds
Feedback
Useful Links
Research Councils
Anniversary
Start your job search with the 3,002 vacancies on jobs.ac.uk...
Search jobs in science, research, academic & related employment in the UK & abroad. Subscribe to Jobs by Email for vacancies in universities, FE colleges, research institutions, commercial & public sector bodies, schools & charities.
Job seekers: select a career sector or type from the lists below to begin your job search. Advertisers: please go here...
Job Search
Search
Search Options: Match all keywords Match any keywords
Search jobs by subject area
Agriculture & Food (62)
Architecture & Planning (82)
Basic Skills (7)
Biological Sciences (487)
Business & Management (442)
Computer Science (356)
Creative Arts & Design (78)
Economics (94)
Education Studies (136)
Engineering & Technology (487)
Environment (132)
Health & Medical (520)
Humanities (95)
Information Management (31)
Languages (86)
Law (41)
Mathematics (197)
Media & Communications (57)
Physical Sciences (468)
Politics & Government (56)
Psychology (128)
Research Methods (22)
Social Sciences & Care (221)
Sport & Leisure (44)
Or search jobs by type / role
Admin, Services & Support (869)
Masters (31)
Premier Adverts
University of Derby Corporate Head Of Corporate Sales
GCHQ Information Specialists
Imperial College London Canterbury Christ Church University Inspire PGCE
University of Surrey Lectureship Theoretical Biological Physics
University of Derby Corporate Head of Learning Solutions
University of Bristol Lecturers in Sociology

Evaluating Sources



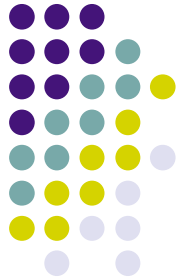
- Author → Who wrote this?
- Date → When? Is it frequently updated?
- Bias → Why?
- Reliability → What else have they written?
- Structure → How?
- URL → Which organisation is behind it

Referencing Sources



- Where appropriate (especially in academic context) you should acknowledge your sources of information
- There are a number of different conventions this, including footnotes and reference lists
- Plagiarism is BAD and a form of intellectual theft!

Key Points



- Internet and Web (HTML, HTTP)
- Browser (FireFox, Explorer, Safari, Opera)
- Search Engines (Google, www.Yahoo)
  - Specific queries
- Classified directories (dir.Yahoo)
  - General queries
- Subject Gateways (BIOME)
  - High quality information
- Evaluate your sources!