

Data Mining and Web Extraction Resources On the Internet

By

Marcus P. Zillman, M.S., A.M.H.A.
Executive Director – Virtual Private Library
zillman@virtualprivatelibrary.com

This April 2008 column **Data Mining and Web Extraction Resources On the Internet** is a comprehensive list of data mining and web extraction sources and sites on the Internet including associated and related online sources. The below list is taken from my current Subject Tracer™ Information Blog titled Data Mining and is constantly updated with Subject Tracer™ bots at the following URL:

<http://www.DataMiningResources.info/>

These resources and sources will help you to discover the many pathways available to you through the Internet for obtaining and locating the resources and sources to help you become part of one of the fastest growing areas on the Internet - Grids and Distributing Computing!.

Data Mining Resources On the Internet:

ACM SIGKDD: Current Explorations Issue

<http://www.acm.org/sigs/sigkdd/explorations/issue.php?issue=current>

Advanced Information Extractor

<http://poorva.com/aie/>

Anthracite Web Mining Desktop Toolkit for MacOS X

<http://www.metafy.com>

Automatic Knowledge Miner - Online data Mining Server

<http://www.auknomi.com/>

AZMY Thinkware -- Data Analysis and Mining Software Tools

<http://www.azmy.com/>

1



April 2008 Zillman Column – Data Mining and Web Extraction Resources On the Internet

<http://www.zillmancolumns.com/>
zillman@VirtualPrivateLibrary.com

© 2008 Marcus P. Zillman, M.S., A.M.H.A.

Bamshad Mobasher Home Page
<http://maya.cs.depaul.edu/~mobasher/index.html>

Benchmarking- Data Mining Benchmarking Association
<http://www.dmbenchmarking.com/>

**Bibliomining for Automated Collection Development in a Digital Library Setting:
Using Data Mining to Discover Web-Based Scholarly Research Works by Dr. Scott
Nicholson**
<http://dlist.sir.arizona.edu/archive/00000625/>
<http://www.BiblioMining.com/>

Biomedical Literature (and text)Mining Publications (BLIMP)
<http://blimp.cs.queensu.ca/>

Bologna Data Mining Centre
http://open.cineca.it/datamining/index_ing.htm

Bot Research
<http://www.BotResearch.info/>

Business Intelligence Data Mining
<http://businessintelligence.ittoolbox.com/nav/t.asp?t=312&p=312&h1=312>

Business Intelligence Text Mining
<http://businessintelligence.ittoolbox.com/nav/t.asp?t=316&p=316&h1=316>

Business Intelligence Resources
<http://www.BIResources.info/>

Business Intelligence Web Mining
<http://businessintelligence.ittoolbox.com/nav/t.asp?t=322&p=322&h1=322>

CCSU - Data Mining
<http://www.ccsu.edu/datamining/>

Center for Automated Learning and Discovery
<http://www.cs.cmu.edu/~cald/>

China Data Mining Research
<http://www.dresearch.net/>



Clementine Data Mining Workbench

<http://www.spss.com/clementine>

CRoss Industry Standard Process for Data Mining (CRISP-DM)

<http://www.crisp-dm.org/>

Current Awareness Discovery Tools on the Internet

<http://zillman.blogspot.com/2004/09/current-awareness-discovery-tools-on.html>

Cypher - Plain Language Access to the Semantic Web

<http://www.monrai.com/products/cypher>

D2K - Data to Knowledge

<http://alg.ncsa.uiuc.edu/do/tools/d2k>

Data Engineering Bulletin

<http://snipurl.com/6o01>

DataFerrett - Data Mining Tool

<http://dataferrett.census.gov/>

Data Fountains: Open Source Internet Resource Discovery and Metadata/Full-Text Generation Service

http://infomine.ucr.edu/Data_Fountains/

Data Mining

<http://zillman.blogspot.com/2005/05/data-mining.html>

Data Mining and Analytical CRM

<http://www.minevision.com/>

Data Mining and Analytic Technologies

<http://www.thearling.com/>

Data Mining and KDD Papers

<http://www.andypryke.com/university/papers.html>

Data Mining and Knowledge Discovery Journal

<http://www.kluweronline.com/issn/1384-5810>



Data Mining - Federal Efforts Cover a Wide Range of Uses Report
<http://www.gao.gov/new.items/d04548.pdf>

DataMiningGrid Consortium
<http://www.datamininggrid.org/>

Data Mining Group (DMG)
<http://www.dmg.org/>

Data Mining, Predictive Modeling, Business Analytics: Training, Consulting & Solutions
<http://www.the-modeling-agency.com/>

Data Mining Research
<http://dmine.blogspot.com/>

Data Mining Resources
<http://www.scd.ucar.edu/hps/GROUPS/dm/dm.html>

Data Mining Resources at CCSU
<http://www.ccsu.edu/datamining/resources.html>

Data Mining Sifts the Gems From Digital Ore By Wilson P. Dizard III
http://www.gcn.com/23_29/news/27420-1.html

Data Mining Software Catalog
<http://www.data-mining-soft.com/>

Data Mining Software - Data Mining Tools - Web Mining Software - Web Mining Tools
<http://www.ir-ware.biz/soften.html#dam>

Data Mining: Text Mining, Visualization and Social Media
http://datamining.typepad.com/data_mining/

Data Mining Using SAS Enterprise Miner by Randall Matignon
<http://www.sasenterpriseminer.com/>

Data Mining, Web Scraping, Web Mining, Data Extraction and Screen Scraping Technology Links
<http://www.connotate.com/>



Data Mining, Web Mining, and Business Intelligence Solutions from Salford Systems

<http://www.salford-systems.com/>

Data Mining, Web Mining and Knowledge Discovery Resources

<http://www.eruditionhome.com/datamining/>

Data Shaping Data Mining Resources

http://www.datashaping.com/data_mining.shtml

Data Sources

<http://www.andypryke.com/university/datasources.html>

DataSpace - Web for Data

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

DbVisualizer - The Universal Database Tool

<http://www.dbvis.com/products/dbvis/>

Deep Web Research

<http://www.DeepWebResearch.info/>

Dig Deep: Data mining Resources for Consultants by Beth Blakely

<http://snipurl.com/6ohs>

DigiCULT Resources - Resource Discovery & Information Retrieval

<http://www.digicult.info/pages/resources.php?t=21>

digitalAGORA

<http://aut.edu/agora/>

Directory of Data Warehouse, Data Mining, and Decision Support Resources

<http://www.infogoal.com/dmc/dmcdwh.htm>

DM Review

<http://www.dmreview.com/>

eBiquity Research Group Blogger

<http://ebiquity.umbc.edu/v2.1/blogger/>



Emping - Data Mining Tool

<http://j-van-thiel.speedlinq.nl/emp/empug.html>

Evolving a Stigmergic Self-Organized Data-Mining by Vitorino Ramos and Ajith Abraham

http://alfa.ist.utl.pt/~cvrm/staff/vramos/ref_50.html

Exclusive Ore, Inc.

<http://www.xore.com/>

Four-T-Nine-R(sm): Data Mining in Web and non-Web Bibliographic Databases

<http://www.public.iastate.edu/~CYBERSTACKS/4T9R.htm>

Genalytics - Advanced Analytics for Marketing and Risk Management

<http://www.genalytics.com/>

GeneMiner

<http://www.kdiscovery.com/>

German Data Mining Portal

<http://mitglied.lycos.de/hpn/DataMining.html>

GMDH - Group Method of Data Handling

<http://come.to/GMDH>

Google Directory - Data Mining

http://directory.google.com/Top/Computers/Software/Databases/Data_Mining/

Graf-FX - Visual Database Data Mining Software

<http://www.gr-fx.com/graf-fx.htm>

Howard D. Wactlar Home Page

<http://www-2.cs.cmu.edu/~hdw/>

Imagination Engines

<http://www.Imagination-Engines.com/>

Information Retrieval (IR) and Information Extraction (IE) on the Web Using Hypertext Meta-Data and Structure

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



InfoVis CyberInfrastructure
<http://iv.slis.indiana.edu/index.html>

Insight Consulting
<http://www.deej.com/insight/>

Integrating Data Mining, Databases and Information Retrieval (IDDI-05)
<http://iddi05.unibg.it/>

International Journal of Business Intelligence and Data Mining (IJBIDM)
<http://www.inderscience.com/ijbidm>

International Journal of Data Mining and Bioinformatics (IJDMB)
<http://www.inderscience.com/ijdmb>

International Journal of Data Warehousing and Mining (IJDWM)
<http://www.igi-global.com/journals/details.asp?id=4291>

KDD-2008
<http://www.kdd2008.com/>

KDnuggets: Data Mining, Web Mining, and Knowledge Discovery Guide
<http://www.kdnuggets.com/>

Knowledge Discovery Resources
<http://www.KnowledgeDiscovery.info/>

Knowledge Discovery Resources 2008 Annotated White Paper Link Compilation by Marcus P. Zillman, M.S., A.M.H.A.
<http://www.KDResources.info/>

KnowleSys - Web Data Extraction
<http://www.knowledsys.com>

LLRX - A Review of TRACFed: Lawyers Strike Gold Mining Government Data
<http://www.llrx.com/features/tracfed.htm>

LLRX - Deep Web Research 2008
<http://zillman.blogspot.com/2007/12/llrx-november-2007-issue-deep-web.html>



Marcus P. Zillman Home Page

<http://www.zillman.us/>

Marti Hearst Home Page

<http://www.sims.berkeley.edu/~hearst/>

MedScan - Automated Scientific Text Mining Tool

<http://www.ariadnegenomics.com/products/medscan.html>

Megaputer - Data Mining, Text Mining and Web Mining Software

<http://www.megaputer.com/>

Metafy Anthracite Web Mining Desktop For MACOS X

<http://www.metafy.com>

Microsoft® Data Mining Project - Efficient Data Exploration and Modeling

<http://research.microsoft.com/dmx/DataMining/>

Mi Li Wo Data Mining Community

<http://www.geocities.com/misforto/>

MonetDB Query Processing at Light Speed

<http://monetdb.cwi.nl/>

National Centre for Text Mining

<http://www.nactem.ac.uk/>

Nesstar

<http://www.nesstar.com/>

NetOwl - Discovery Software from SRA International

<http://www.netowl.com/index.html>

NewsTin - Multilingual News Search

<http://www.newstin.com/>

Oceanstore Project

<http://oceanstore.cs.berkeley.edu/>

OntoMiner: Bootstrapping and Populating Ontologies From Domain Specific Web Sites

<http://www.public.asu.edu/~hdavulcu/VLDB-WS03.pdf>

Open Access Now - Data Mining Open Access Research

<http://snipurl.com/6o2r>

Open Directory Project - Data Mining

http://www.dmoz.org/Computers/Software/Databases/Data_Mining/

Orange - Data Mining Software

<http://www.ailab.si/orange>

PC AI Magazine

<http://www.pcai.com/pcai>

Pentaho BI Project - Open Source Business Intelligence

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

PEPITe S.A. - Unlock Your Knowledge

<http://www.pepите.be/>

Prediction Markets

<http://www.PredictionMarkets.com/>

Predictive Model Markup Language (PMML)- SourceForge.net: Project Info

<http://sourceforge.net/projects/pmml>

Predictive Model Markup Language (PMML)

<http://www.oasis-open.org/cover/pmml.html>

PubGene™ Database and Tools

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

QL2 Software - Unstructured Data Management and Web Mining Software

<http://www.ql2.com/>

Raghu Ramakrishnan Home Page

<http://www.cs.wisc.edu/~raghu/>



RapidMiner - Open Source Data Mining Tool

<http://rapid-i.com/content/blogcategory/10/69/>

reSearcher

<http://researcher.sfu.ca/>

Rexer Analytics - Data Mining and CRM Analytics

<http://www.rexeranalytics.com/>

Ron Kohavi Home Page

<http://robotics.stanford.edu/~ronnyk/>

SAS - Data and Text Mining

<http://snipurl.com/6033>

Sciengy RPF!™

<http://sciengy.com/products/>

ScrapeGoat

<http://www.scrapegoat.com/>

Screen-Scraper - Extracts Information From Web Sites

<http://www.screen-scraper.com/>

Searching the Internet

<http://www.SearchingTheInternet.info/>

SIGKDD2008

<http://www.acm.org/sigs/sigkdd/kdd2008/>

SIGKDD - ACM Special Interest Group - Knowledge Discovery in Data and Data Mining

<http://www.acm.org/sigs/sigkdd/>

SpagoBI - Unified Open Source Platform for Business Intelligence

<http://www.spagoworld.org/ecm/faces/public/guest/home/solutions/spagobi>

SPIDER: Scalable, Parallel and Interactive Data Mining and Exploration at Rensselaer

<http://www.cs.rpi.edu/~zaki/datamining.html>



Special Interest Group - Knowledge Discovery in Data and Data Mining - SIGKDD Explorations Newsletter

<http://www.acm.org/sigs/sigkdd/explorations/>

SPSS, Data Mining, Statistical Analysis Software, Predictive Analysis, Predictive Analytics, Decision Support Systems

<http://www.netgen.com/>

SQL Server Data Mining

<http://www.sqlserverdatamining.com/>

StarProbe Visual Data Miner

<http://www.roselladb.com/starprobe.htm>

Statistical Data Mining Tutorials - Tutorial Slides by Andrew Moore

<http://www-2.cs.cmu.edu/~awm/tutorials/index.html>

Statoo Consulting Data Mining Resources Links

<http://www.statoo.com/en/resources/anthill/Datamining/>

Streamload - Share Videos and Photos - Online MP3 Storage and Access

<http://www.streamload.com/>

Survey of DHS Data Mining Activities - Office of Information Technology

<http://snipurl.com/wgff>

T2K - Text to Knowledge

<http://alg.ncsa.uiuc.edu/do/tools/t2k>

Text Data Mining

<http://www.sims.berkeley.edu/~hearst/talks/dm-talk/>

Text Mining, Web Mining, Information Retrieval and Extraction from the WWW References

http://filebox.vt.edu/users/wfan/text_mining.html

The Data Mine

<http://www.the-data-mine.com/>

The Data Warehouse

<http://www.datawarehouse.com/>



theinfo.org - For People With Large Data Sets

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

The National Center for Data Mining

<http://www.ncdm.uic.edu/>

The National Centre for Text Mining: Aims and Objectives by Sophia Ananiadou, Julia Chruszcz, John Keane, John McNaught and Paul Watry

<http://www.ariadne.ac.uk/issue42/ananiadou/>

The Seventh Asia Pacific Web Conference (APWeb05)

<http://apweb05.csm.vu.edu.au/index.asp>

Topic Detection and Tracking (TDT)

<http://www.itl.nist.gov/iaui/894.01/tests/tdt/>

Unit Miner - Web Data Extraction Software

<http://www.qualityunit.com/unitminer/web-extraction-tool.html>

Using the Internet As a Dynamic Resource Tool for Knowledge Discovery

<http://zillman.blogspot.com/2004/09/using-internet-as-dynamic-resource.html>

Vendor-Neutral Public Courses on Data Mining Strategy, Methods & Practice

<http://www.the-modeling-agency.com/training>

VisitorVille - Web Site Intelligence

<http://www.visitorville.com/>

Visual Analytics

<http://www.visualanalytics.com/>

Web Curator Tool (WCT)- Management of Selective Web Harvesting Process

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

Web Data Extractors - White Paper Link Compilation

<http://zillman.blogspot.com/2004/09/web-data-extractors.html>

Web Data Mining

http://www.blossom.com/web_mining.html

Web Data Mining Online Papers

<http://maya.cs.depaul.edu/~classes/cs589/papers.html>

WebFarming.com - Farming the Web for Systematic Business Intelligence

<http://www.webfarming.com/>

Web-Harvest - Open Source Web Data Extraction Tool

<http://web-harvest.sourceforge.net/index.php>

Web Harvesting by Russell Kay

<http://snipurl.com/78sy>

Web Mining

<http://galeas.de/webmining.html>

Web Mining

<http://www.cs.umbc.edu/~ajoshi/web-mine/>

Web Mining - Business Intelligence

<http://snipurl.com/6ogb>

WebQL

<http://www.ig.com.au/WebQL.htm>

Web Search and Mining

<http://intelligent-web.org/wsm/>

Weka 3 - Data Mining with Open Source Machine Learning Software in Java

<http://www.cs.waikato.ac.nz/~ml/weka/index.html>

White Papers by Marcus P. Zillman, M.S., A.M.H.A.

<http://www.WhitePapers.us/>

WizSoft - Data and Text Mining

<http://www.wizsoft.com/>

Yahoo Groups - Data Mining

<http://groups.yahoo.com/group/datamining2/>

<http://groups.yahoo.com/group/datamining/>



Web Extraction Resources On the Internet:

Advanced Information Extractor (AIE)

<http://poorva.com/aie/>

Anomic HTTP Proxy

<http://freshmeat.net/releases/173068/>

Anthracite

<http://freshmeat.net/projects/anthracite/>

Automated RSS Scraper Scripts

<http://www.djeaux.com/rss/>

Automated Information Solutions

<http://www.automated-info-solutions.com/>

Automatic Information Extraction From Semi-Structured Web Pages By Pattern Discovery

<http://portal.acm.org/citation.cfm?id=640423&dl=ACM&coll=portal>

Beautiful Soup

<http://freshmeat.net/projects/beautifulsoup/>

BotMakers

<http://www.BotMakers.com/>

Bot Research

<http://www.BotResearch.info/>

BYU Data Extraction Research Group

<http://www.deg.byu.edu/>

Captiva Software: Digital Information Capture Software

<http://www.captivasoftware.com/index.asp>

Client-Side Deep Web Data Extraction

<http://www.tic.udc.es/~mad/publications/ceceast2004.pdf>

Compaq's Web Language

<http://research.compaq.com/SRC/WebL/>

14



April 2008 Zillman Column – Data Mining and Web Extraction Resources On the Internet

<http://www.zillmancolumns.com/>
zillman@VirtualPrivateLibrary.com

© 2008 Marcus P. Zillman, M.S., A.M.H.A.

Dapper - Extract and Use Website Information for Mashups

<http://www.dappit.com/index.php>

Data Extractors and Pass-Through Systems – A Selected List

<http://www.chass.utoronto.ca/datalib/misc/dli/extracts.htm>

Data Extractors – Special Report

<http://snipurl.com/8flw>

Data Extraction Web Scraping and Web Mining Using Screen Scraping Technology

<http://www.connotate.com/>

Data Locators on the Web

http://www.nd.edu/~lsrteach/archive_old%20site/01207all/presentation_final.htm

Data Mining Resources

<http://www.DataMiningResources.info/>

Deep Web Research

<http://www.DeepWebResearch.info/>

Effective Web Data Extraction with Standard XML Technologies

<http://www10.org/cdrom/papers/102/>

ExtractData Technologies - SearchExtract Software

<http://www.extradata.com/>

Extracting Knowledge

http://www.intelligentkm.com/feature/010507/feat1.jhtml?_requestid=185742

FerretSoft

<http://www.FerretSoft.com/>

Ficstar Software - Web Data Extraction

<http://www.ficstar.com/index.html>

Google™ Directory – Extractors

<http://directory.google.com/Top/Computers/Software/Shareware/Windows/Internet/Email/Extractors/>



Imagination Engines

<http://www.Imagination-Engines.com/>

Information Foraging and Extraction Techniques for Internet-Based Literature and Data

http://www.asceditor.usm.edu/ASC%202006%20CD/2006pro/2006/CEGE03_Hannon06_7900.htm

Information Retrieval (IR) and Information Extraction (IE) on the Web

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

iOpus® Internet Macros™

<http://www.iopus.com/iim.htm>

iWebMiner – Web Data Mining and Content Scraping

<http://www.iwebminer.com/>

Kapow RoboSuite Platform - Solutions for Data Collection

http://www.kapowtech.com/solutions_datacollection.htm

Knowledge Discovery Resources

<http://www.KnowledgeDiscovery.info/>

Knowlesys® - Web Data Extraction, Web Grabber and Screen Scraper

<http://www.knowlesys.com/index.htm>

Lingway

<http://www.lingway.fr/en/>

Mastering Data Extraction

<http://www.dbmsmag.com/9606d05.html>

Metadata Extraction Tool

<http://meta-extractor.sourceforge.net/>

mnot: xpath2rss - HTML->RSS scraper

<http://www.mnot.net/xpath2rss/>

NewsClipper.com - Snip and Ship Dynamic News Content to Your Web Pages

<http://www.newsclipper.com/>



NQL Technologies

<http://www.nqltech.com/>

Pervasive Data Management and Integration Products

<http://www.pervasive.com/>

QL2 Software - Unstructured Data Management and Web Mining Software

<http://www.ql2.com/>

REBOL Technologies

<http://www.rebol.com/>

ScrapeForge

<http://freshmeat.net/projects/scrapeforge/>

ScrapeGoat

<http://www.ScrapeGoat.com/>

Scraper

<http://freshmeat.net/projects/scraper/>

Screen-Scraper

<http://freshmeat.net/projects/screenscraper/>

Screen-Scraper – Extracts Information From Web Sites

<http://www.Screen-Scraper.com/>

Screenscraping the Senate by Paul Ford

<http://www.xml.com/pub/a/2004/09/01/hack-congress.html>

Screen Snarfs - Darkspell: Perl Code for Screen Scrapers

<http://www.darkspell.com/gadgets/snarfs/>

Search and Replace with TextPipe Pattern Matching

<http://www.crystalsoftware.com.au/textpipe.html>

Sitescooper: Scoop Websites Onto Your PalmPilot

<http://sitescooper.org/>

Text Mining and Web-Based Information Retrieval Reference

http://filebox.vt.edu/users/wfan/text_mining.html



That Robot Site – Smart Internet Solutions

<http://www.thatrobotsite.com/>

Unit Miner - Web Data Extraction Software

<http://www.qualityunit.com/unitminer/web-extraction-tool.html>

URL Link Extractor

<http://www.fuddyduddy.connectfree.co.uk/urlgen.htm>

Visual Web Spider

<http://www.newprosoft.com/>

Visual Web Task

<http://www.lencom.com/VisualWTSite.html>

W3C Publishes Data Extraction Language (DEL) as W3C Note

<http://xml.coverpages.org/ni2001-11-06-a.html>

Web Data Extraction Software

<http://www.tethyssolutions.com/web-data-extraction.htm>

Web Data Extractor

<http://www.webextractor.com/>

Web Data Extractor

<http://www.rafasoft.com/>

Web Data Mining

http://www.blossom.com/web_mining.html

Web-Harvest – Open Source Web Data Extraction Tool

<http://web-harvest.sourceforge.net/index.php>

Web Mining and Unstructured Data Management Solutions – QL2 Software

<http://www.ql2.com/>

WebQL

<http://www.ig.com.au>

WebScraper Plus +

<http://www.velocityscape.com/>



Website Extractor

<http://www.hot-shareware.com/internet-tools/website-extractor/>

Website Extractor – Offline Browser

<http://www.internet-soft.com/extractor.htm>

Website Scraping

<http://www.websitescraping.com/>

Web Spider, Link Extraction, And Other Extractor Products

<http://www.pjltechnology.com/>

WebWrappers - Agents for E-Cataloging, Automation, Data Aggregation and Integration

<http://www.webwrappers.com/>

wisosoft - Intelligent Agents for Web Aggregation and Deep Web Processing

<http://www.oejhw.or.at/rand.htm>

WWW::Extractor

<http://freshmeat.net/projects/ade/>

XRay Web Scraping Tool

<http://freshmeat.net/projects/xrayguibasedwebscrapingtool/>

Subject Tracer™ Information Blogs

Subject Tracer™ Information Blogs created and developed by the Virtual Private Library™ combine the best of the latest tools on the Internet. Using bots, blogs and news aggregators the Subject Tracer™ Information blogs generate RSS feeds with the latest resources to create a current information resource flow through niched subject tracers. I am proud to be the creator of the Internet's first Subject Tracer™ Information Blogs:

Virtual Private Library™

<http://www.VirtualPrivateLibrary.com/>

Agriculture Resources

<http://www.AgricultureResources.info/>



Artificial Intelligence Resources

<http://www.AIResources.info/>

Astronomy Resources

<http://www.AstronomyResources.info/>

Auction Resources

<http://www.AuctionResources.info/>

Biological Informatics

<http://www.BiologicalInformatics.info/>

Biotechnology Resources

<http://www.BiotechnologyResources.info/>

Bot Research

<http://www.BotResearch.info/>

Business Intelligence Resources

<http://www.BIResources.info/>

ChatterBots

<http://www.ChatterBots.info/>

Data Mining Resources

<http://www.DataMiningResources.info/>

Deep Web Research

<http://www.DeepWebResearch.info/>

Directory Resources

<http://www.DirectoryResources.info/>

eCommerce Resources

<http://eCommerceResources.info/>

Elder Resources

<http://www.ElderResources.info/>

Employment Resources

<http://www.EmploymentResources.info/>



Entrepreneurial Resources
<http://www.EntrepreneurialResources.info/>

Financial Sources
<http://www.FinancialSources.info/>

Finding People
<http://www.FindingPeople.info/>

Games Resources
<http://www.GamesResources.info/>

Genealogy Resources
<http://www.GenealogyResources.info/>

Grant Resources
<http://www.GrantResources.info/>

Grid Resources
<http://www.GridResources.info/>

Healthcare Resources
<http://www.HealthcareResources.info/>

Information Futures Markets
<http://www.InformationFutureMarkets.com/>

Information Quality Resources
<http://www.InformationQualityResources.info/>

Internet Alerts
<http://www.InternetAlerts.info/>

Internet Demographics
<http://www.InternetDemographics.info/>

Internet Experts
<http://www.InternetExperts.info/>

Internet Hoaxes
<http://www.InternetHoaxes.info/>



Journalism Resources
<http://www.JournalismResources.info/>

Knowledge Discovery
<http://www.KnowledgeDiscovery.info/>

Military Resources
<http://www.MilitaryResources.info/>

Outsourcing/Offshoring Information and Resources
<http://www.OutsourcingOffshore.us/>

Privacy Resources
<http://www.PrivacyResources.info/>

Reference Resources
<http://www.ReferenceResources.info/>

Research Resources
<http://www.ResearchResources.info/>

RestStress™
<http://www.RestStress.com/>

Script Resources
<http://www.WcriptResources.info/>

ShoppingBots
<http://www.ShoppingBots.info/>

Social Informatics
<http://www.SocialInformatics.info/>

Statistics Resources
<http://www.StatisticsResources.info/>

Student Research
<http://www.StudentResearch.info/>

Theology Resources
<http://www.TheologyResources.info/>



Tutorial Resources

<http://www.TutorialResources.info/>

World Wide Web Reference

<http://www.WWWReference.info/>

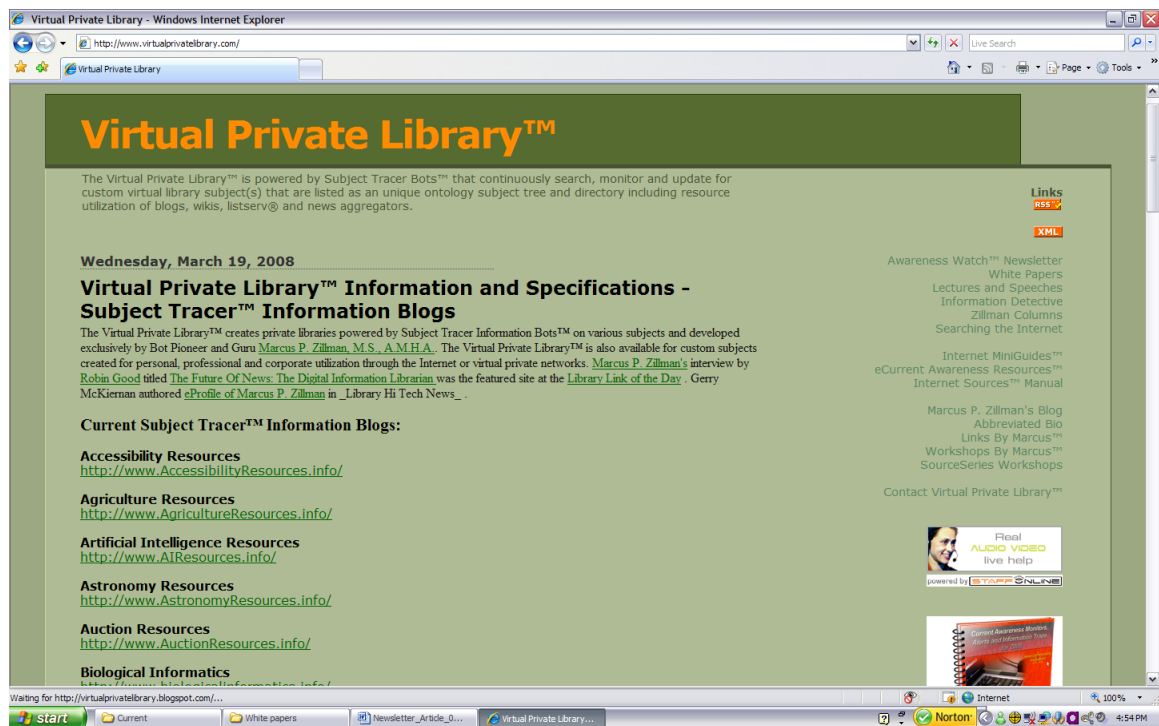


Figure 1: Virtual Private Library™

Author Information: Marcus P. Zillman, M.S., A.M.H.A. Executive Director of the Virtual Private Library is an international Internet expert, author, keynote speaker and corporate consultant in the area of information retrieval, knowledge discovery, knowledge harvesting, artificial intelligence and bots/intelligent agents. He has created numerous world wide web sites including 47 Subject Tracer™ Information Portals and Blogs; written a number of internet miniguides, white papers, manuals and books; hosted over 160 weekly Internet television shows, writes a weekly and monthly column on Current Awareness on the Internet; writes a monthly newsletter Awareness Watch and delivers keynote presentations throughout the international marketplace. He also actively delivers one and two day workshops for key industry sectors displaying how the Internet can be used as a tool to maintain current awareness and professional competencies.

23



April 2008 Zillman Column – Data Mining and Web Extraction Resources On the Internet

<http://www.zillmancolumns.com/>
zillman@VirtualPrivateLibrary.com

© 2008 Marcus P. Zillman, M.S., A.M.H.A.

Additional websites by Marcus P. Zillman, M.S., A.M.H.A.:

Marcus P. Zillman's Blog

<http://www.zillman.us/>

Marcus P. Zillman Abbreviated Bio

<http://www.zillman.info/>

White Papers by Marcus P. Zillman

<http://www.WhitePapers.us/>

Internet MiniGuides™

<http://www.InternetMiniguide.com/>

Awareness Watch™ Newsletter

<http://www.AwarenessWatch.com/>

Marcus P. Zillman's Columns

<http://www.ZillmanColumns.com>

Current Awareness Monitors, Alerts and Information Traps for 2008

<http://www.eCurrentAwareness.com/>

Internet Sources™ Manual

<http://www.InternetSources.info/>

Links By Marcus™

<http://www.LinksByMarcus.com/>

Workshops By Marcus™

<http://www.WorkshopsByMarcus.com/>

SourceSeries Internet Research Workshops

<http://www.SourceSeries.com/>

Watch Marcus™

<http://www.WatchMarcus.com/>

listen to marcus™

<http://www.ListenToMarcus.com>



Research White Papers, Articles, Lectures and Speeches by Marcus P. Zillman, M.S., A.M.H.A.:

Academic and Scholar Search Engines and Sources

<http://zillman.blogspot.com/2004/12/academic-and-scholar-search-engines.html>

Bots, Blogs and News Aggregators

<http://www.BotsBlogs.com/>

Business Intelligence Online Resources

<http://www.BIOnlineResources.info/>

Current Awareness Discovery Tools on the Internet

<http://zillman.blogspot.com/2004/09/current-awareness-discovery-tools-on.html>

Deep Web Research 2008 Article - LLRX

<http://zillman.blogspot.com/2007/12/llrx-november-2007-issue-deep-web.html>

Healthcare Bots and Subject Directories

<http://www.HealthcareBots.info/>

Knowledge Discovery Resources 2008

<http://www.KDResources.info/>

Lectures and Speeches by Marcus P. Zillman, M.S., A.M.H.A.

<http://snipurl.com/57jp>

Online Research Browsers

<http://zillman.blogspot.com/2004/10/online-research-browsers-internet.html>

Online Research Tools

<http://zillman.blogspot.com/2004/09/online-research-tools.html>

Online Social Networking

<http://zillman.blogspot.com/2004/09/online-social-networking-internet.html>

Searching the Internet

<http://www.SearchingTheInternet.info/>

Using the Internet As a Dynamic Resource Tool for Knowledge Discovery

<http://zillman.blogspot.com/2004/09/using-internet-as-dynamic-resource.html>



Web Data Extractors

<http://zillman.blogspot.com/2004/09/web-data-extractors.html>

White Papers By Marcus P. Zillman, M.S., A.M.H.A.

<http://www.WhitePapers.us/>

Internet Tutor by Marcus P. Zillman, M.S., A.M.H.A.

<http://www.InternetTutor.info/>

Visit this site to learn about the availability of Marcus P. Zillman to tutor you or your associate one on one in the privacy of your residence or office on the latest happenings of the Internet including Internet basics to advanced Internet searching using bots and creating your own personal blog

Internet Speaking by Marcus P. Zillman, M.S., A.M.H.A.

<http://www.InternetSpeaker.net>

Visit this site to learn about Marcus P. Zillman's speaking engagements for your organization meetings and events. View and listen to his previous presentations as well as his weekly television shows

Internet Consulting by Marcus P. Zillman, M.S., A.M.H.A.

<http://InternetConsultant.BlogSpot.com/>

Visit this site to obtain information about obtaining the consultation services of Marcus P. Zillman for your company including eCommerce audits, utilization of bots, blogs and news aggregators or the creation of your own personal virtual private library powered by Subject Tracer™ Information bots!

Internet Sources™ Manual

<http://www.InternetSources.info>

Marcus P. Zillman's latest 378 page manual **Internet Sources™** is now available for purchase online and for immediate download. This book makes a great reference resource for the "newbie" to the Internet as well as the seasoned veteran "Internaut".

Current Awareness Monitors, Alerts and Information Traps for 2008

<http://www.ecurrentAwareness.com/>

Marcus P. Zillman's latest report Current Awareness Monitors, Alerts and Information Traps for 2008 is now available for purchase online and for immediate download. This report is a comprehensive listing of the latest resources, sources and sites for current awareness on the Internet. This is a must read for anyone who must stay current in their profession and/or business activity as the list of URLs will keep you at the leading edge of your career.



Market Intelligence Resources 2008

<http://www.MarketIntelligenceResources.com/>

Marcus P. Zillman's just released professional Internet MiniGuide is titled Market Intelligence Resources 2008 and is now available for purchase online and immediate download. This 110 page digital minguide represents a comprehensive listing of the latest resources, sources and sites to discover the latest Market Intelligence sources available on the Internet with many of them freely available! Designed specifically for today's entrepreneur, professional and/or investor.

