

Data Mining Resources on the Internet

By

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

Data Mining Resources on the Internet is a comprehensive listing of data mining resources currently available on the Internet. The below list of sources is taken from my Subject Tracer™ Information Blog titled Data Mining Resources 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 to find the latest data mining sources and sites.

Data Mining Resources:

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



Data Mining Resources White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

© 2005 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/>

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.dmrsearch.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>

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>

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, 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/>

DataSpace - Web for Data

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

Data Sources

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



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/>

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/>

KDD-2006

<http://www.kdd2006.com/>

KDnuggets: Data Mining, Web Mining, and Knowledge Discovery Guide

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

Knowledge Discovery Resources

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

Knowledge Discovery Resources 2005 Annotated White Paper Link Compilation by Marcus P. Zillman, M.S., A.M.H.A.

<http://zillman.blogspot.com/2005/03/knowledge-discovery-resources-2005.html>



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 2005

<http://zillman.blogspot.com/2004/09/deep-web-research-llrx-featured.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/>

Nesstar

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

NetOwl - Discovery Software from SRA International

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

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>

PEPITe S.A. - Unlock Your Knowledge

<http://www.pepite.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/>

reSearcher

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



Ron Kohavi Home Page

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

SAS - Data and Text Mining

<http://snipurl.com/6o33>

Sciengy RPF!™

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

Searching the Internet

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

SIGKDD2006

<http://www.acm.org/sigs/sigkdd/kdd2006/>

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

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

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/>

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/>

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/>

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 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 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/>

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/>

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/>

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/>



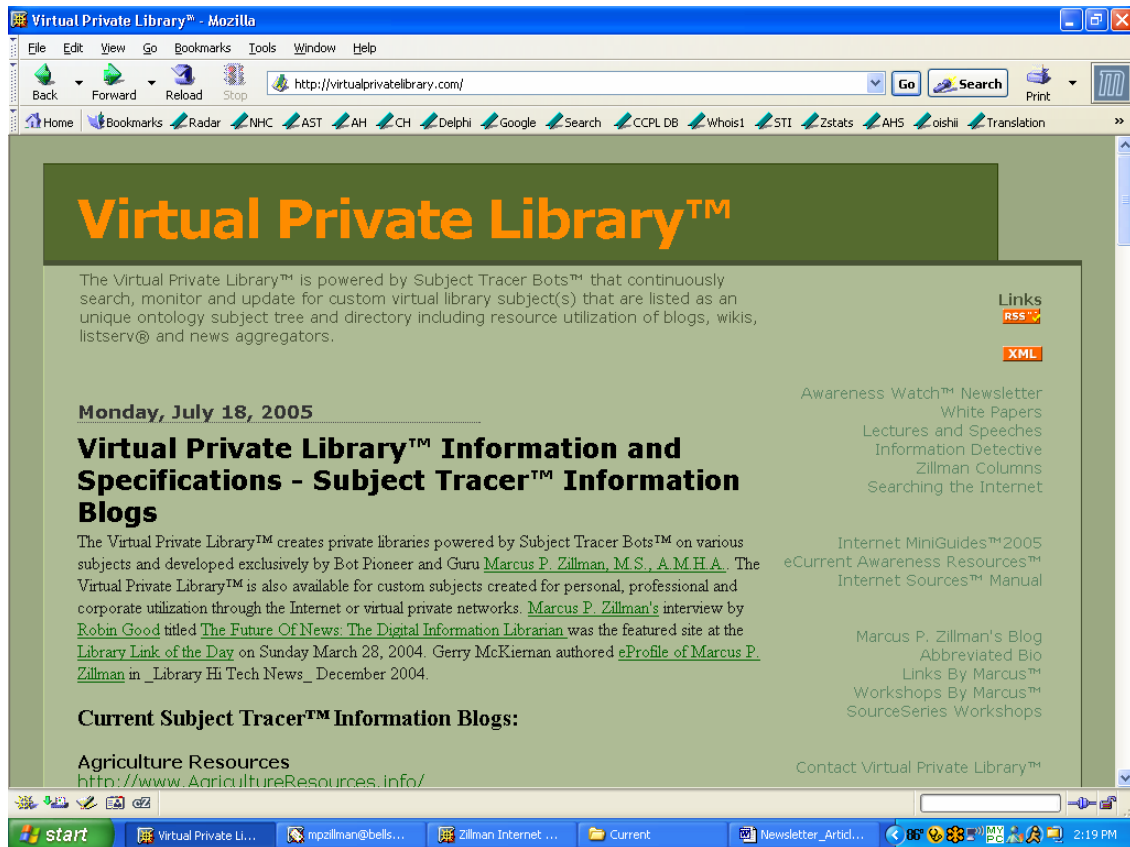


Figure 2 Virtual Private Library™

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 45 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. 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://marcuszillman.blogspot.com/>

White Papers by Marcus P. Zillman
<http://www.WhitePapers.us/>

Internet MiniGuides™ 2005
<http://www.InternetMiniguide.com/>

Awareness Watch™ Newsletter
<http://www.AwarenessWatch.com/>

Marcus P. Zillman's Columns
<http://www.ZillmanColumns.com>

eCurrent Awareness Resources™ 2005 Business Intelligence Report
<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://zillman.blogspot.com/2005/04/business-intelligence-online-resources.html>

Current Awareness Discovery Tools on the Internet

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

Deep Web Research 2005 Article - LLRX

<http://zillman.blogspot.com/2004/09/deep-web-research-llrx-featured.html>

Healthcare Bots and Subject Directories

<http://zillman.blogspot.com/2005/05/healthcare-bots-and-subject.html>

Information Detective – Online Streaming Tutorial Videos

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

Knowledge Discovery Resources 2005

<http://zillman.blogspot.com/2005/03/knowledge-discovery-resources-2005.html>

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!

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". Visit the following site for additional information and online ordering fulfillment:

Internet Sources™ Manual
<http://www.InternetSources.info>

Marcus P. Zillman's latest report eCurrent Awareness Resources 2005 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. Visit the following site for additional information and online ordering fulfillment:

eCurrent Awareness Resources 2005
<http://www.ecurrentAwareness.com/>

