

Data Mining Resources on the Internet 2019

By

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

Data Mining Resources on the Internet 2019 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.

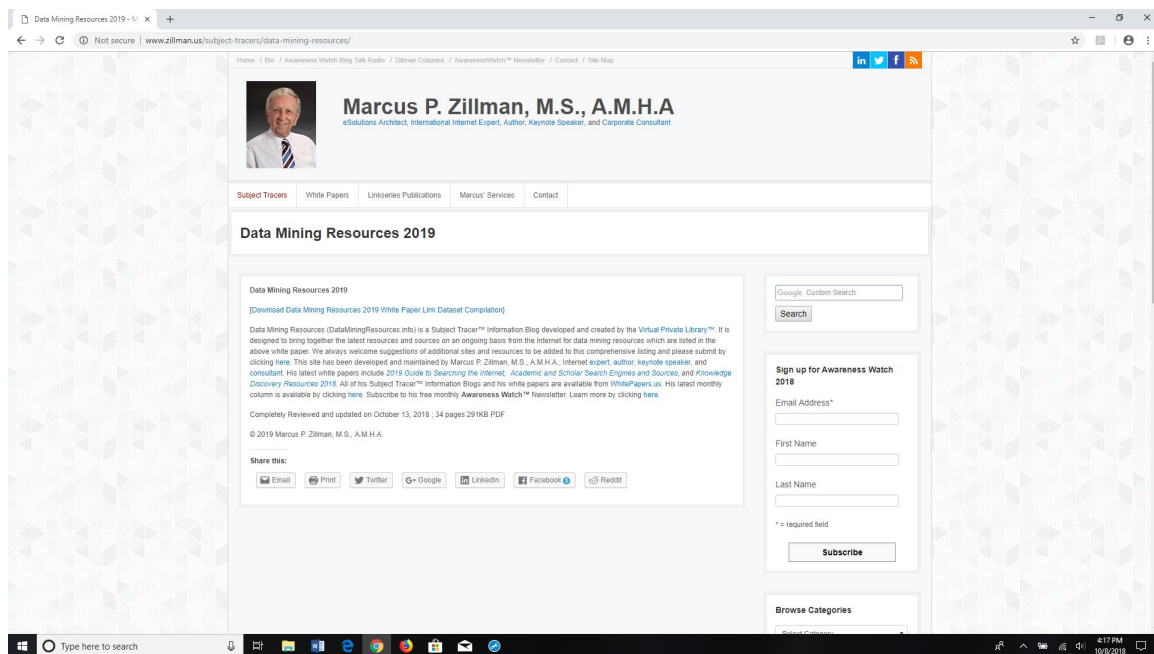


Figure 1: Data Mining Resources 2019 Subject Tracer™ Information Blog

1



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

<http://www.DataMiningResources.info/>
zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

© 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019 Marcus P. Zillman, M.S., A.M.H.A.

Data Mining Resources 2019:

45 Great Resources for Learning Data Mining Concepts and Techniques

<https://www.import.io/post/38-great-resources-for-learning-data-mining-concepts-and-techniques/>

50 Data Mining Resources – Tutorials, Techniques and More

<http://www.ngdata.com/data-mining-resources/>

80legs – Custom Web Crawlers for Crawling and Processing Web Content

<http://www.80legs.com/>

Advanced Analytics – Unstructured Data Mining

<https://advisory.kpmg.us/deal-advisory/data-driven-tech/advanced-analytics.html>

An Evaluation of Data Mining Methods and Tools

<http://www.idi.ntnu.no/~dingsoyr/project/report.html>

An Overview of Data Mining in Road Traffic and Accident Analysis

http://www.jcaksrce.org/upload/49121180_vol2i4p6.pdf

ACM SIGKDD: Current Explorations Issue

<http://www.kdd.org/explorations/issue.php?issue=current>

Analytics, Data mining and Data Science

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

Apache Pig – Platform for Analyzing Large Datasets

<http://pig.apache.org/>

Applications of Modern Heuristics and Data Mining Techniques

http://www.people.vcu.edu/~mmanic/papers/grads/McCarty_08_MHandAdvancDMTechniqs.pdf

Artificial Neural Networks for Data Mining

http://www.iasri.res.in/sscnars/data_mining/4-Artificial%20Neural%20Networks_Amrender.pdf



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

ARTstor - Digital Image Library for Education and Scholarship
<http://www.artstor.org/>

Azure: Data Mining in the Cloud
<http://sqlmag.com/reporting/azure-data-mining-cloud>

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

Best Data Mining Tools – 2018 Reviews, Pricing and Demos
<https://www.softwareadvice.com/bi/data-mining-comparison/>

BI-DW - Business Intelligence and Data Warehousing Directory
<http://www.bi-dw.info/>

Big Data Analytics with Oracle Advanced Analytics
https://blogs.oracle.com/datamining/entry/big_data_analytics_with_oracle

Big Oil Goes Mining for Big Data
<http://www.technologyreview.com/news/427876/big-oil-goes-mining-for-big-data/>

BioHunter _ Research Articles Finder
<https://www.biohunter.in/>

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

Bot Research 2019
<http://www.BotResearch.info/>

Bringing Data Mining Into The Mainstream New York Times Article
http://bits.blogs.nytimes.com/2010/07/26/bringing-data-mining-into-the-mainstream/?_r=0

Business Intelligence Resources 2018
<http://www.BIResources.info/>

Calculating Costs of a Data Mining System
<http://www.eweek.com/c/a/Data-Storage/Calculating-Costs-of-a-DataMining-System>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

CCSU - Data Mining

<http://web.ccsu.edu/datamining/>

Center for Automated Learning and Discovery

<http://www.ml.cmu.edu/>

Cogitum Co-Citer

<http://www.cogitum.com/co-tracker-text/more.shtml>

Contentmine - Text and Data Mining Open Source Tools

<http://contentmine.org/>

Copyright Clearance Center

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

COREMINE Medical – Biomedical Mindmap

<http://www.coremine.com/medical/>

Current Awareness Discovery Tools on the Internet 2019

<http://www.zillman.us/white-papers/current-awareness-discovery-tools-on-the-internet/>

DataMelt - Computation and Visualization Environment

<http://jwork.org/dmelt/>

Data Mining 101 Tools and Techniques

<https://iaonline.theiia.org/data-mining-101-tools-and-techniques>

Data Mining Courses on Campus

<http://www.nytimes.com/2012/07/22/education/edlife/colleges-awakening-to-the-opportunities-of-data-mining.html>

Data Mining Tutorial

http://www.tutorialspoint.com/data_mining/dm_evaluation.htm

Data Engineering Bulletin

http://tab.computer.org/tcde/bull_about.html

DataFerrett - Data Mining Tool

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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

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

Data Mining Amazon Web Services (AWS) Big Data

<https://aws.amazon.com/big-data/>

Data Mining and Analytic Technologies

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

Data Mining and KDD Papers

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

Data Mining and Knowledge Discovery Journal

<http://link.springer.com/journal/10618>

Data Mining and Predictive Analytics

<http://abbottanalytics.blogspot.com/>

Data Mining Applications in Transportation Engineering

<http://www.slideshare.net/Tommy96/data-mining-applications-in-transportation-engineering>

Data Mining Blog EMC

<http://bigdatablog.emc.com/tag/data-mining/>

Data Mining Case Study - Data based storytelling for marketing

<https://automatedinsights.com/blog/making-data-driven-storytelling-easy-for-marketers/>

Data Mining Case Study - Mining complex financial information

<https://automatedinsights.com/blog/how-financial-services-companies-can-win-over-millennials-better-customer-communication-through-automation/>

Data Mining Certification from IBM

<http://www-03.ibm.com/certify/certs/47100402.shtml>

Data Mining Concepts

<https://msdn.microsoft.com/en-us/library/ms174949.aspx>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Data Mining ebook: Theories Algorithms and Examples

<https://www.routledge.com/products/9781439808382>

Data Mining for the Masses

<http://www.onlineprogrammingbooks.com/data-mining-masses/>

Data Mining - Federal Efforts Cover a Wide Range of Uses Report

<http://www.gao.gov/new.items/d04548.pdf>

Data Mining Glossary

<http://www.gartner.com/it-glossary/data-mining/>

Data Mining Group (DMG)

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

Data Mining in Banking and its Applications

<http://www.scribd.com/doc/270947349/Data-Mining-Banking#scribd>

Data Mining Managed Data Mining Services Teradata

<http://www.teradata.com/Search.aspx?id=8550&q=data%20mining&LangType=16393&LangSelect=true>

Data Mining Oil and Gas Hydrocarbon Exploration Data

<http://www.analytics-magazine.org/november-december-2011/695-how-big-data-is-changing-the-oil-a-gas-industry>

Data Mining OLAP Cubes model

http://zulfiqar.typepad.com/zulfiqars_web/data_mining/

Data Mining, Predictive Modeling, Business Analytics: Training, Consulting & Solutions

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

Data Mining Primer from Oracle

http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm

Data Mining Publications from Google

<https://research.google.com/pubs/DataMining.html>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Data Mining Resources 2019

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

Data Mining Resources

<http://www.cs.purdue.edu/homes/ayg/CS590D/resources.html>

Data Mining Resources

<http://datamining.togaware.com/>

Data Mining Resources

<http://datamining.togaware.com/>

Data Mining Resources at CCSU

<http://web.ccsu.edu/datamining/resources.html>

Data Mining Solutions for Oil & Gas

<http://optelligentsolutions.com/Solutions.html>

Data Mining Table Analysis Tool

[https://technet.microsoft.com/en-us/library/dd299414\(v=sql.100\).aspx](https://technet.microsoft.com/en-us/library/dd299414(v=sql.100).aspx)

Data Mining Techiques in CRM

<http://www.data-miners.com/>

Data Mining: Technology and Policy The DHS Privacy Office

http://www.dhs.gov/xlibrary/assets/privacy/privacy_rpt_datamining_200812.pdf

Data Mining: Text Mining, Visualization and Social Media

http://datamining.typepad.com/data_mining/

Data Mining Tools

<http://www.icsti.org/IMG/pdf/VTTDataMiningTools.pdf>

Data Mining Tools

<http://thenewstack.io/six-of-the-best-open-source-data-mining-tools/>

**Data Mining, Web Scraping, Web Mining, Data Extraction and Screen Scraping
Technology Links**

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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

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

Data Mining White Paper – Free Best Practices Guide

<http://www.sas.com/data-mining/>

Data Mining White Paper from Intel - Turning Big Data Into Big Insights

<http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/big-data-visualization-turning-big-data-into-big-insights.pdf>

Data Mining – Wikipedia

https://en.wikipedia.org/wiki/Data_mining

Datanami – Big Data, Big Analytics, and Big Insights

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

Data-PASS

<http://www.data-pass.org/>

Data Science Toolkit

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

Datasets for Data Mining and Data Science

<https://www.kdnuggets.com/datasets/index.html>

Data Shaping Data Mining Resources

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

Data Sources

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

Data Visualizations Derived From Data Mining Big Data

<http://exploringdata.github.io/>

Data Warehousing and Data Mining

<http://www.dei.unipd.it/~capri/SI/MATERIALE/DWDM0405.pdf>

Data Warehousing and Data Mining

http://www.inf.unibz.it/dis/teaching/DWDM/slides2012/dw1_introduction.pdf



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

DbVisualizer - The Universal Database Tool

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

DeepDive - Analyze Data On a Deeper Level Than Ever Before

<http://deepdive.stanford.edu/>

Deep Learning for Java - Open Source, Distributed, Deep Learning Library for the JVM

<https://deeplearning4j.org/>

Deep Web Research and Discovery Resources 2018

<http://DeepWeb.us/>

Digital Operating Systems Tools and Resources 2019

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

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

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

DiscoverText - Capture Text Data and Crunch Your Data

<http://discovertext.com/>

Distributed Data Mining in Credit Card Fraud Detection

<http://cs.fit.edu/~pkc/papers/ieee-is99.pdf>

Easy Data Mining Software

<https://www.tableau.com/>

Easy PDF Cloud

<https://www.easypdfcloud.com/>

eBiquity Research Group Blogger

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

Early Canadiana Online

<http://www.canadiana.ca/>

Elastic Web Mining Talk

<http://www.slideshare.net/kkrugler/elastic-web-mining-2407818>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

ELKI: Environment for Developing KDD-Applications Supported by Index-Structures

<https://elki-project.github.io/>

EU Open Data Portal

<https://data.europa.eu/euodp/en/home>

Everything You Wanted to Know About Data Mining but Were Afraid to Ask by Alexander Furnas

<http://www.theatlantic.com/technology/archive/2012/04/everything-you-wanted-to-know-about-data-mining-but-were-afraid-to-ask/255388/>

Exclusive Ore, Inc.

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

GeneMiner

<http://www.biomedcentral.com/1471-2105/8/S8/P3>

Google BigQuery query cloud based datasets

<https://cloud.google.com/bigquery/>

Google Refine 2.0 – Power Tool for Data Wranglers

<https://github.com/OpenRefine>

Great War Primary Documents Archive

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

Healthdata.gov

<https://www.healthdata.gov/>

History of Data Mining by Raymond Li

<http://rayli.net/blog/data/history-of-data-mining/>

Howard D. Wactlar Home Page

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

IBM Data Mining Cognos Business Solutions

<http://www.cognoise.com/index.php?topic=19929.0>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Imagination Engines

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

Indiegogo Datasets

<https://webrobots.io/indiegogo-dataset/>

Infochimps.com

<http://www.infochimps.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>

International Journal of Business Intelligence and Data Mining (IJBIDM)

<http://www.inderscience.com/jhome.php?jcode=ijbidm>

International Journal of Data Mining and Bioinformatics (IJDMB)

<http://www.inderscience.com/jhome.php?jcode=ijdmb>

International Journal of Data Warehousing and Mining (IJDWM)

<http://www.igi-global.com/journal/international-journal-data-warehousing-mining/1085>

Internet Archive

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

Inter-university Consortium for Political and Social Research (ICPSR)

<http://www.icpsr.umich.edu/>

InvestigateIX Search Engine and Text Mining Toolbox

<https://www.mandalka.name/investigateix/>

Jaspersoft® ETL - The Open Source Data Integration Platform

<https://community.jaspersoft.com/project/jaspersoft-etl>

Journal of Data Mining and Knowledge Discovery

<http://www.bioinfo.in/contents.php?id=42&page=aim>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Junar – The Open Data Platform

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

Kaggle – Go from Big Data to Big Analytics

<http://kaggle.com/>

KDD-2008

<http://www.kdd2008.com/>

KDD-2009

<http://www.kdd.org/conferences/kdd-2009-paris-france-june-28-july-1>

KDD-2010

<http://www.kdd.org/conferences/kdd-2010-washington-dc-july-25-28>

KDD-2011

<http://www.kdd.org/conferences/kdd-2011-san-diego-ca-august-21-24-2011>

KDD-2012

<http://www.kdd.org/conferences/kdd-2012-august-12-16-2012-beijing-china>

KDD-2014

<http://www.kdd.org/kdd2014/>

KDD-2015

<http://www.kdd.org/kdd2015/>

KDD-2016

<http://www.kdd.org/kdd2016/>

KDD-2017

<http://www.kdd.org/kdd2017/>

KDD-2018

<http://www.kdd.org/kdd2018/>

KDD-2019

<http://www.kdd.org/kdd2019/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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

KEEL (Knowledge Extraction based on Evolutionary Learning)
<http://www.keel.es/>

Kickstarter Datasets
<https://webrobots.io/kickstarter-datasets/>

KNIME – Konstanz Information Miner Open Source Software
<http://www.knime.org/>

Knowledge Discovery Resources 2018
<http://www.KnowledgeDiscovery.info/>

Knowledge Discovery Resources 2018 Annotated White Paper Link Compilation by Marcus P. Zillman, M.S., A.M.H.A.
<http://www.KDResources.info/>

Knowledge Enterprise Semantic Intelligence Suite
<http://transinsight.com/>

KnowleSys - Web Public Opinion Monitoring
<http://www.knowledsys.com>

LingPipe - Information Extraction and Data Mining Tools
<http://alias-i.com/lingpipe/>

LoginWorks - Advanced Solutions - Data Mining and Web Scraping
<http://www.loginworks.com/>

Machine Learning from Scratch
<https://github.com/eriklindernoren/ML-From-Scratch>

Mallet - MACHine Learning for Language Toolkit
<http://mallet.cs.umass.edu/>

Marriott Library at the University of Utah Digital Collections
<http://www.lib.utah.edu/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Marti Hearst Home Page

<http://people.ischool.berkeley.edu/~hearst/>

Media Patterns - Detecting Patterns in the Global Media Content

<http://mediapatterns.enm.bris.ac.uk/>

Megaputer - Data Mining and Text Mining Software

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

Microsoft® Data Mining Project - Efficient Data Exploration and Modeling

<http://research.microsoft.com/en-us/projects/datamining/>

Minerazzi – Your Search-and-Mine Ecosystem

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

Mining Road Traffic Accident Data

<http://ai-d.org/pdfs/Beshah.pdf>

Mining Spatial Data of Traffic Accidents

<http://www.stat-d.si/mz/mz5.1/lavrac.pdf>

MIT OpenCourseWare study and certification Data Mining Discipline

<http://ocw.mit.edu/courses/sloan-school-of-management/15-062-data-mining-spring-2003/>

MOA (Massive Online Analysis)

<http://moa.cms.waikato.ac.nz/>

MoData - Big Data Resources

<http://www.mo-data.com/>

MonetDB Query Processing at Light Speed

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

Mozenda – Data Extraction and Comprehensive Web Data Gathering

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

National Archives, London

<http://nationalarchives.gov.uk/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

National Centre for Text Mining (NaCTeM)

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

National Science Digital Library (NSDL)

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

National Technical Information Service (NTIS)

<http://www.ntis.gov/>

Neural Networks in Data Mining

<http://www.jatit.org/volumes/research-papers/Vol5No1/1Vol5No6.pdf>

Nesstar – Publish Data on the Web

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

NetOwl – Entity Extraction and Entity Analytics for Big Data

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

New York Public Library

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

Nuix - eDiscovery and Electronic Investigation Software

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

Oceanstore Project

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

Online News Archive

<https://onlinenewsarchive.com/>

OntoMiner: Bootstrapping and Populating Ontologies From Domain Specific Web Sites

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

Open Data Barometer

<http://www.opendataresearch.org/project/2013/odb>

Open Data Handbook - Guides, Case Studies and Resources for Government and Civil Society On the What, Why and How of Open Data

<http://opendatahandbook.org/>

15



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

© 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019 Marcus P. Zillman, M.S., A.M.H.A.

Open Data Inception

<http://opendatainception.io/>

Open Data Institute

<https://theodi.org/>

Open Data Inventory (ODIN)

<http://odin.opendatawatch.com/>

Open Data Network

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

Open Datasets

<https://github.com/caesar0301/awesome-public-datasets>

<https://www.kaggle.com/datasets>

<https://www.data.gov/>

<https://www.quora.com/Where-can-I-find-large-datasets-open-to-the-public>

<https://aws.amazon.com/public-datasets/>

<https://data.world/>

<http://data.worldbank.org/>

<https://public.enigma.com/>

<http://repository.upenn.edu/mead>

<http://catalog.data.gov/dataset>

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

<http://publicdata.eu/>

Open Educational Resources (OER) Sources 2019

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

Open Educational Resources (OER): Statistics

<https://guides.ou.edu/OER/statistics>

OpenMinted - Open Service Oriented e-Infrastructure for Scientific and Scholarly Text and Data Mining

<http://openminted.eu/>

Open/Public Data Sources

<http://www.scaleunlimited.com/datasets/public-datasets/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Open Source Data Mining Tools

<http://www.scaleunlimited.com/oss/open-source-data-mining-tools/>

Open Source Data Warehousing, Big Data Analytics

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

Oracle Data Mining

<http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/overview/index.html>

Oracle Knowledge base about Big Data Mining

<http://docs.oracle.com/apps/search/search.jsp?word=data+mining&product=b28359-01&book=b28129>

Orange - Open Source Data Visualization and Analysis for Novice and Experts

<http://orange.biolab.si/>

Overview - Open Source Document Mining

<https://blog.overviewdocs.com/>

PC AI Magazine Artificial Intelligence

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

Pentaho BI Project - Open Source Business Intelligence

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

PEPITe S.A. - Unlock Your Knowledge

<http://www.pepite.be/>

Prediction Markets 2018

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

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

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

Predictive Model Markup Language (PMML)

<http://xml.coverpages.org/pmml.html>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Probabilistic Data Models for Web Analytics and Data Mining

<https://highlyscalable.wordpress.com/2012/05/01/probabilistic-structures-web-analytics-data-mining/>

Proxycrawl - Stay Anonymous While Crawling the Web

<https://proxycrawl.com/>

PubChase – Discover Biomedical Research of Interest To You

<https://www.pubchase.com/>

QDA Miner Lite (Freeware)

<http://provalisresearch.com/products/qualitative-data-analysis-software/freeware/>

QL2 Software - Unstructured Data Management and Web Mining Software

<http://www.q12.com/>

QueryTree - Explore Data Without Code

<http://querytreeapp.com/>

Raghu Ramakrishnan Home Page

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

RapidMiner - Open Source Data Mining Tool

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

Rattle - Data Mining Toolkit in R

<https://code.google.com/p/rattle/>

re3data.org - 2,000 Data Repositories

<https://www.re3data.org/>

Recommended Books on Data Mining

http://www.albionresearch.com/books/data_mining.php

reSearcher

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

Rexer Analytics – Analytic and CRM Consulting

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Ron Kohavi Home Page

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

SAS - Data and Text Mining

<http://www.sas.com/technologies/analytics/datamining/index.html>

SAS What is Data Mining

http://www.sas.com/en_us/insights/analytics/data-mining.html

SCaVis - Scientific Computation and Visualization Environment

<http://jwork.org/scavis/>

Scientific Data Repository - Real Time Visualization and Exploration Techniques

<http://www.mlvis.com/platform.php>

Screen-Scraper – Data Extraction Software and Services

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

Searching the Internet 2018 – The Primer – 2019 Guide to Searching the Internet

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

Semantic Scholar - Free Scientific Literature Search and Discovery

<http://allenai.org/semantic-scholar/>

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

<https://en.wikipedia.org/wiki/SIGKDD>

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

Slideshare Presentations About Data Mining - a List

<http://www.kdnuggets.com/2014/11/most-popular-slideshare-presentations-data-mining.html>

Slideshare Presentation about Data Mining

<http://www.slideshare.net/smj/data-mining-slides>

Smithsonian/NASA Astrophysics Data System (ADS)

<http://adsabs.harvard.edu/index.html>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Snorkel: A System for Fast Training Data Creation

<https://hazyresearch.github.io/snorkel/>

Social Buzz Bot 2019 – Business Intelligence Data Mining for Information Discovery from Social Communities [PDF file download]

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

Software Suites for Data Mining, Analytics, and Knowledge Discovery

<http://www.kdnuggets.com/software/suites.html>

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

Explorations Newsletter

<http://www.kdd.org/explorations/>

SPMF - Open Source Data Mining Library

<http://www.philippe-fournier-viger.com/spmf/>

Stanford Data Mining Course cs345a course handouts

<http://web.stanford.edu/class/cs345a/handouts.html>

SQL Server Data Mining

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

Statistical Analysis and Data Mining

<http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291932-1872>

Statistical Data Mining Tutorials - Tutorial Slides by Andrew Moore

<http://www.autonlab.org/tutorials/index.html>

Statistics Resources and Big Data 2018

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

Statoo Statistical Consulting + Data Analysis + Data Mining

<http://www.statoo.com/en/>

Streaming Data Mining

http://www.cs.yale.edu/homes/el327/papers/streaming_data_mining.pdf

Talend Open Data Solutions

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

20



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

© 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019 Marcus P. Zillman, M.S., A.M.H.A.

Tanagra Project - Free Data Mining Software for Academic and Research Purposes
<http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html>

Texifter - Search, Sift, Sort & Classify Documents
<http://texifter.com/>

Text Data Mining
<http://people.ischool.berkeley.edu/~hearst/talks/dm-talk/>

Text Mining for Scholarly Communications and Repositories
<http://www.nactem.ac.uk/tm-ukoln.php>

The Archaeology Data Service (ADS)
<http://archaeologydataservice.ac.uk/>

The Centre for Contemporary Canadian Art - Canadian Art Database Project
<http://ccca.concordia.ca/>

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

The History Data Service (HDS)
<http://hds.essex.ac.uk/>

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 New York Times Article Search API
<http://developer.nytimes.com/>

The Open Access Digital Library
<http://grweb.coalliance.org/oadl/oadl.html>

The Ultimate Artificial Intelligence Resources Guide by Kyle Poyar
<http://labs.openviewpartners.com/artificial-intelligence-resources-guide/>

Togaware - Data Mining Resources
<http://datamining.togaware.com/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

T-Rex (Trainable Relation Extraction)

<http://sourceforge.net/projects/t-rex/>

Truthy - Analyze and Visualize the Diffusion of Information on Twitter

<http://truthy.indiana.edu/>

Try Data Mining Queries Interactively Online using sample dataset

<http://overpass-turbo.eu/>

UC Irvine Machine Learning Repository

<https://archive.ics.uci.edu/ml/index.php>

Udemy Course About Data Mining

<https://www.udemy.com/data-mining/>

University of Florida Digital Collections (UFDC)

<http://ufdc.ufl.edu/>

University of North Texas Digital Collections

<http://digital.library.unt.edu/explore/collections/>

Using the Internet As a Dynamic Resource Tool for Knowledge Discovery 2018

<http://www.zillman.us/white-papers/using-the-internet-as-a-dynamic-resource-tool-for-knowledge-discovery/>

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

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

VentureSource - Global Database on Companies Backed by Venture Capital and Private Equity

<https://www.dowjones.com/products/pevc/>

VisitorVille - Web Site Intelligence

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

Wallmine - Wall Street Data Mining

<https://wallmine.com/>

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

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Web Data Extractors 2018 - White Paper Link Compilation

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

Web-Harvest - Open Source Web Data Extraction Tool written in Java

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

Web Harvesting by Russell Kay

http://www.computerworld.com/s/article/93919/Web_Harvesting?taxonomyId=062

Webhose.io - Turn Unstructured Web Content Into Machine-Readable Data Feeds That You Can Consume On Demand

<https://webhose.io/>

Weka 3: Data Mining Software in Java

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

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/>

World Bank Datasets For Data Mining

<http://data.worldbank.org/data-catalog/research-datasets-analytical-tools>

YouTube Analytics and Data Mining

<http://www.nextanalytics.com/wp-content/uploads/2015/01/how-to-data-mine-analyze-youtube-excel-macro-addin.pdf>

Zentut - What is Data Mining Tutorial

<http://www.zentut.com/data-mining/what-is-data-mining/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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/>

Accessibility Resources

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

Agriculture Resources

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

AnswerSpot

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

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/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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/>

Education and Academic Resources

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

Elder Resources

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

Employment Resources

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

Entrepreneurial Resources

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

Fact Checkers Directory

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

Financial Sources

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

Finding People

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Games Resources

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

Genealogy Resources

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

Grant Resources

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

Green Files

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

Grid, Distributed and Cloud Computing Resources

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

Healthcare Resources

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

Information Futures Markets

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

Information Quality Resources

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

International Trade Resources

<http://www.InternationalTradeResources.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/>



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Intrapreneurial Resources

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

Journalism Resources

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

Knowledge Discovery

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

Military Resources

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

New Economy Analytics, Resources and Alerts

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

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.ScriptResources.info/>

ShoppingBots

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

Social Informatics

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Statistics Resources and Big Data
<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/>

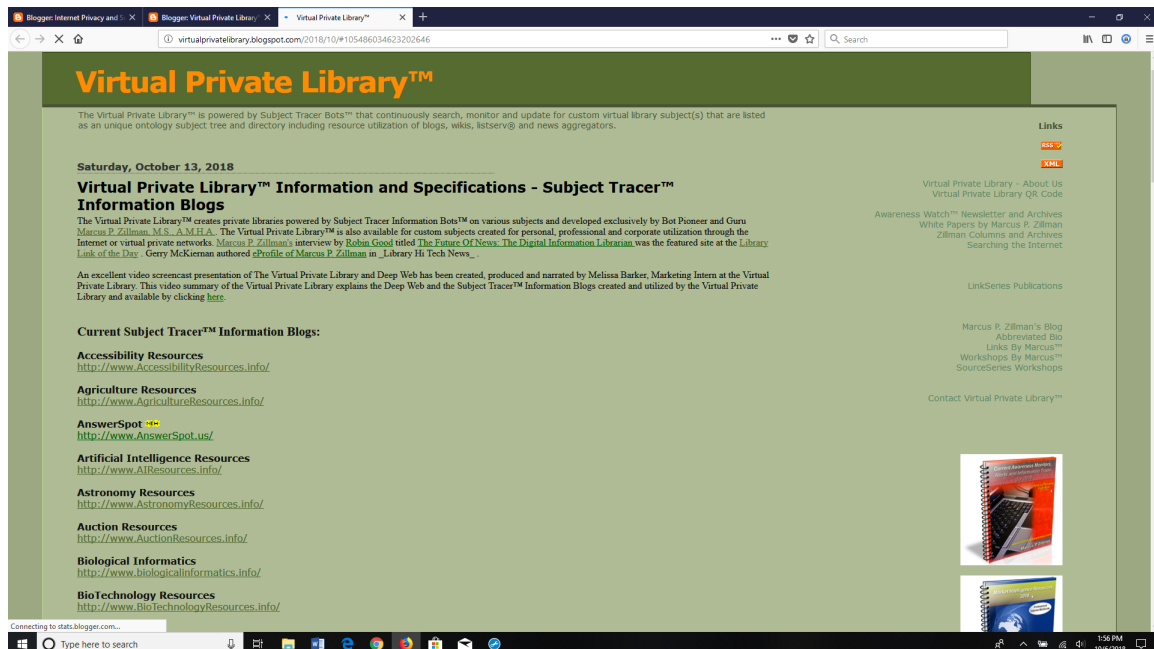


Figure 2: Virtual Private Library™

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 54 Subject Tracer™ Information Portals and



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

<http://www.DataMiningResources.info/>
zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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 (20,000+ Postings)

<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>

LinkSeries Publications

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

Links By Marcus™

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

Workshops By Marcus™

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

SourceSeries Internet Research Workshops

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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

2018 Directory of Directories

<http://www.2018DirectoryOfDirectories.com/>

Academic and Scholar Search Engines and Sources

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

Bots, Blogs and News Aggregators 2019

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

Business Intelligence Online Resources 2018

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

Cloud Computing Resources Primer 2018

<http://www.zillman.us/white-papers/grid-distributed-and-cloud-computing-resources-primer/>

Current Awareness Discovery Tools on the Internet 2019

<http://www.zillman.us/white-papers/current-awareness-discovery-tools-on-the-internet/>

Deep Web Research and Discovery Resources 2018 Online White Paper

<http://zillman.blogspot.com/2017/01/llrx-deep-web-research-and-discovery.html>

<http://DeepWeb.us/>

eMarketing MiniGuide 2018

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

eReference Library Link Toolkit 2018

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

Finding Experts By Using the Internet 2018

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

Finding People Resources and Sites 2019

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Healthcare Bots and Subject Directories 2018

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

Knowledge Discovery Resources 2018

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

New Economy Resources 2018

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

Online Research Browsers 2018

<http://www.zillman.us/white-papers/online-research-browsers/>

Online Research Tools

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

Online Social Networking 2018

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

Searching the Internet 2018 – The Primer – 2019 Guide to Searching the Internet

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

Using the Internet As a Dynamic Resource Tool for Knowledge Discovery 2018

<http://www.zillman.us/white-papers/using-the-internet-as-a-dynamic-resource-tool-for-knowledge-discovery/>

Web Data Extractors 2018

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

Web Guide for the New Economy 2018

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

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

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



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

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!

Current Awareness Monitors, Alerts and Information Traps

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

Marcus P. Zillman's latest report Current Awareness Monitors, Alerts and Information Traps is 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

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

Marcus P. Zillman's just released professional Internet MiniGuide is titled Market Intelligence Resources and is available for purchase online and immediate download. This 193 page digital miniguide 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.



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462

Entrepreneurial Links 101

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

Marcus P. Zillman's newly released 231 page eReference digital book for the up and coming entrepreneur. Entrepreneurial Links 101 gives an alphabetical listing of the very best Internet and World Wide Web sites covering Entrepreneur Resources, Business Intelligence Resources and an extremely comprehensive list of Online Research Tools. This is considered by many to be the entrepreneur's bible for finding relevant and competent online resources!

Internet Privacy and Security Resources

<http://www.InternetPrivacySecurity.net/>

Marcus P. Zillman's latest eReference digital publication is a selected comprehensive alphabetical listing of the latest resources and sites covering all aspects of privacy and security currently available over the Internet. From the board room to the family room, these resources and sites give you the information you need to maintain your privacy and security as you use the Internet in your business and personal life.

Research Resources Online Guide

<http://www.ResearchResourcesOnline.net/>

Marcus P. Zillman's latest [LinkSeries Publication](#) is a 340 page digital guide of a selected comprehensive alphabetical listing of the latest and greatest resources and sites covering all areas of research that is currently available over the Internet. The guide covers online research resources and tools for the Newbie to research as well as the Seasoned researcher. Contents include: a) Research Resources, b) Research Tools, c) Student Research Resources Toolkit, d) Knowledge Discovery/Management and Data Mining Resources, e) Knowledge Discovery/Retrieval and the World Wide Web Resources, and f) Subject Tracer™ Information Blogs.

The Survivor's Manual for The New Economy.

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

Marcus P. Zillman's latest LinkSeries Publication is a 239 page digital read that gives excellent resources and annotated sources for the new economy analytics, alerts, ecommerce, financial sources, invisible and deep web resources, social and business networking sources along with new economy competitive and business intelligence resources and an extremely comprehensive listing of new economy online tools.



[Released October 13, 2018]

Data Mining Resources 2019 White Paper Link Compilation

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

zillman@VirtualPrivateLibrary.com

eVoice: 800-858-1462