

# Web Data Extractors 2021

By

**Marcus P. Zillman, M.S., A.M.H.A.**  
**Executive Director - Virtual Private Library**  
<http://www.VirtualPrivateLibrary.org>

The November 2020 Zillman Column features **Web Data Extractors 2021** and is a very comprehensive listing of Internet and Web data extractor resources and sites on the Internet for the latest competent web data extraction markets and research. The below list of sources is taken partially from my Subject Tracer™ white paper titled Web Data Extractors 2021 and is updated monthly with Subject Tracer™ bots at the following URL:

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

These resources will help you to discover the many pathways available to you through the Internet to find the latest web data extractor sources and sites. These resources will be used to find required web data extractor information for the very latest research and discovery resources and happenings to find that next data extraction in the [New Normal!!](#)

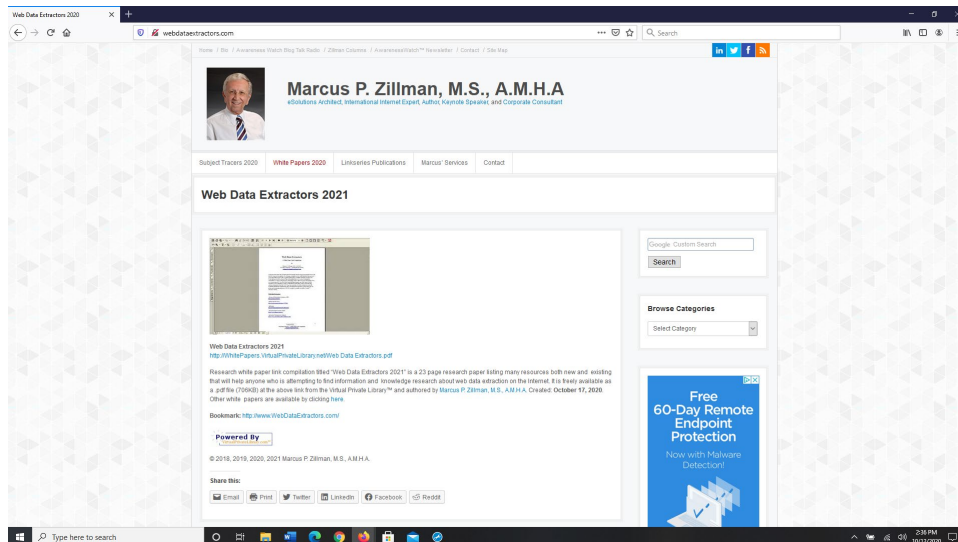


Figure 1: Web Data Extractors 2021



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

## **Web Data Extractors 2021:**

80legs - Powerful and Economical Service Platform for Crawling and Processing Web Content

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

Agenty – Hosted Web Scraping Tool

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

Anthracite

<http://freecode.com/projects/anthracite>

Apify – Web Scraping Platform for Coders

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

Aristo - Answer Questions with a Knowledgeable Machine

<http://allenai.org/aristo/>

artoo.js - The Client-Side Scraping Companion

<http://medialab.github.io/artoo/>

AutoMate - Automate Data Extraction

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

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://freecode.com/projects/beautifulsoup>

Beautiful Soup - HTML/XML Parser for Quick Turnaround Screen Scraping and Web Data Extraction

<http://www.crummy.com/software/BeautifulSoup/>



November 2020 Zillman Column – Web Data Extractors 2021

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

[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

bli solutions Weather Predictive Analytics  
<http://www.biasolutions.com/>

Bot Research 2020  
<http://www.BotResearch.info/>

BYU Data Extraction Research Group  
<http://www.deg.byu.edu/>

Captiva Software: OpenText Captiva – Capture and Transform Content  
<http://www.emc.com/enterprise-content-management/captiva/captiva.htm>

Client-Side Deep Web Data Extraction  
<http://www.tic.udc.es/~mad/publications/ceceast2004.pdf>

CloudScape – Extract, Enrich and Connect  
<http://www.cloudscape.com/>

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

Common Crawl  
<http://www.commoncrawl.org/>

Connotate – Web Data Extraction and Monitoring  
<http://www.connotate.com/>

ContextMiner - Tools to Collect Data, Metadata and Contextual Information  
<http://www.contextminer.org/>

cQuery - Content Query Engine  
<http://cquery.com/>

CrawlMonster  
<http://www.crawlmonster.com/>

Crawly  
<http://crawly.diffbot.com/>

Create a Crawler - Extract Data From an Entire Website  
<https://www.import.io/>



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

cURL groks URLs - Command Line Tool for Transferring Data  
<http://curl.haxx.se/>

Data Extraction Services  
<http://www.dataextractionservices.com/>

DataHen – Advanced Web Scraping and Data Extraction Services  
<https://www.datahen.com/>

Data Mining Resources 2021  
<http://www.DataMiningResources.info/>

Data Miner – Extract Data From any Website in Seconds  
<https://data-miner.io/>

Dataminr - Real-time Information Discovery  
<http://www.dataminr.com/>

Data Scraper – Easy Web Scraping with Google Chrome  
<https://chrome.google.com/webstore/detail/data-scraper-easy-web-scr/nndknejnldbepjfgmncbggmopgden?hl=en-US>

DataSift - Powerful Social Data Platform  
<http://datasift.com/>

Data Toolbar – Web Data Extraction Software Made Simple  
<http://datatoolbar.com/>

DataWatch Monarch – Self-Service Data Preparation  
<http://www.datawatch.com/>

DataWrangler - Data Cleaning and Transformation Tool  
<http://vis.stanford.edu/wrangler/>

Deep Web Research 2020  
<http://www.DeepWebResearch.info/>

DEiXTo – Powerful Web Data Extraction Tool Based on W3C DOM  
<http://deixto.com/>



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

dexi.io – Web Data Processing for Professionals – Extract, Enrich and Connect  
<https://dexi.io/>

DiffBot AI – Web Data Extraction Using Artificial Intelligence  
<http://www.DiffBot.com/>

Diggernaut - Data Scraping - Turn Website Content Into Datasets  
<https://www.diggernaut.com/>

Digital Footprints - Collect Facebook Data  
<http://digitalfootprints.dk/>

DiscoverText - Import, Sort, Distribute and Analyze Electronic Content from eMail,  
Document Repositories, and Social Media  
<http://discovertext.com/>

Easy PDF Cloud  
<https://www.easypdfcloud.com/>

Easy Web Extract – Best Tool for Web Scraping  
<http://webextract.net/>

eGrabber - Data Capture Tools  
<http://www.egrabber.com/>

Facepager - Fetching Public Data From Facebook  
<https://github.com/strohne/Facepager>

Ficstar Software - Web Data Extraction  
<http://www.ficstar.com/>

File Information Tool Set (FITS)  
<https://projects.iq.harvard.edu/fits>

FMiner – Web Scraping Software  
<http://www.fminer.com/>

Fresh WebSuction  
<http://www.freshwebmaster.com/>



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Get Automated Data Extraction

<https://info.helpsystems.com/am-data-extraction-and-movement/>

GetData.io - Get Valuable Data from the Web in 3 Steps

<https://getdata.io/>

Grepsr – Web Scraping Made Simple, Fast and Manageable

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

Harkive – Data Collection – Multiple Sources/Single Database

<http://harkive.org/datcolzap/>

Helium Scraper

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

How to Scrape Data from a Website Using Python

<https://www.codementor.io/oluwagbengajoloko/how-to-scrape-data-from-a-website-using-python-n3fmtc63q>

How To Use A Data-Scaping Tool to Extract Data From Webpages

<https://www.maketecheasier.com/use-data-scraping-tool-extract-data-from-web-pages>

Huginn - Your Agents Are Standing By

<https://github.com/cantino/huginn>

Hunter - Connect With Anyone

<https://hunter.io/>

HYPHE - Web Corpus Curation Tool Featuring A Research-Driven Web Crawler

<http://hyphe.medialab.sciences-po.fr/>

iCyte - Your Research Anywhere

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

iMacros – Data Extraction

<http://imacros.net/overview>

Imagination Engines

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



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Import.io - Turn the Web Into Data With Extractors, Crawlers and Connectors  
<https://import.io/>

InfoExtractor - Extracts Relevant Information from Blogs, YouTube and Twitter  
<http://www.infoextractor.org/>

Information Retrieval (IR) and Information Extraction (IE) on the Web  
<http://www.webir.org/>

Instaloader – Download Pictures or Videos and Metadata from Instagram  
<https://instaloader.github.io/>

Instamancer – Scrape Instagram’s API with Puppeteer  
<https://adamsm.com/instamancer/>

Introduction to Information Retrieval  
<http://www-nlp.stanford.edu/IR-book/>

Introduction to Web Scraping Using Python  
<https://github.com/qut-dmrc/web-scraping-intro-workshop>

Ion – On-demand IP Rotation for Faster Web Scraping  
<https://www.anonymizer.com/>

iRobotSoft – Visual Web Scraping and Web Automation  
<http://irobotsoft.com/>

iWeb Scraping Services  
<http://www.iwebscraping.com/>

Jaspersoft® ETL - The Open Source Data Integration Platform  
<https://community.jaspersoft.com/project/jaspersoft-etl>

JetOctopus – Crawler for Big Web Sites  
<https://jetoctopus.com/>

Junar - Discovering Data  
<http://www.junar.com/>

Karma - Data Integration Tool  
<http://www.isi.edu/integration/karma/>



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Knowledge Discovery Resources 2020  
<http://www.KnowledgeDiscovery.info/>

Knowlesys® - Web Data Extraction, Web Grabber and Screen Scraper  
<http://www.knowlesys.com/index.htm>

Liberty Metrics – Web Scraping Services  
<http://libertymetrics.com/>

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

Listly - Fully Automated Web Scraping Service  
<http://listly.io/>

LoginWorks – On Demand Webpage Scraping Services  
<https://www.loginworks.com/>

Marquee – Professional Web Scraping Services  
<https://marqueedata.com/>

Metadata Extraction Tool  
<http://meta-extractor.sourceforge.net/>

Mozenda – Comprehensive Web Data Gathering  
<http://www.mozenda.com/>

Netlytic - Making Sense of Online Conversations  
<https://netlytic.org/home/>

Newprosoft – Web Data Extraction Software  
<http://newprosoft.com/>

Octo – Expose Data From Any Database As Web Service  
[https://octoproject.github.io/octo-cli/?utm\\_campaign=Data\\_Elixir&utm\\_source=Data\\_Elixir\\_303](https://octoproject.github.io/octo-cli/?utm_campaign=Data_Elixir&utm_source=Data_Elixir_303)

Octoparse – Automated Web Scraping Software  
<http://www.octoparse.com/>



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.



## Open Datasets

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

<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/>

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

## OutWit Hub - Harvest the Web With Your Own Web Collection Engine

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

## ParseHub – Web Crawling Using Machine Learning

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

## Pervasive Data Management and Integration Products

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

## Priceonomics - Crawl Data From the Web

<http://priceonomics.com/>

## Proxycrawl - Stay Anonymous While Crawling the Web

<https://proxycrawl.com/>

## QL2 Software - Unstructured Data Management and Web Mining Software

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

## Quick Code

<https://quickcode.io/>

## re3data.org - 2,000 Data Repositories

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

## REBOL Technologies

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

## ReVerb - Open Information Extraction Software

<http://reverb.cs.washington.edu/>



November 2020 Zillman Column – Web Data Extractors 2021

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

[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

ScrapeForge

<http://freecode.com/projects/scrapeforge>

Semantic Scholar - Free Scientific Literature Search and Discovery

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

Sequentum – Unlock the World’s Largest Data Source

<https://sequentum.com/>

ScrapeHero

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

ScrapeStorm – AI-Powered Visual Web Data Extractor

<https://www.scrapestorm.com>

Scraper

<http://freecode.com/projects/scraper>

ScrapingBot – API You Need for Efficient Scraping

<https://www.scraping-bot.io/>

ScrapingHub – Cloud Based Data Extraction Tool

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

Scraping Solutions – When the Solution You Seek Seems Impossible

<https://www.scrapingsolutions.com.au/>

Scrapy – Open Source Web Scraping Framework for Python

<http://scrapy.org/>

Screen-Scraper

<http://freecode.com/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>

Search and Replace with TextPipe Pattern Matching

<http://www.datamystic.com/textpipe.html>



November 2020 Zillman Column – Web Data Extractors 2021

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

[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Sensible Code

<http://sensiblecode.io/>

Simple Scraper – Extract Data From Any Website in Seconds

<https://simplescraper.io/>

Social Media Data Collection Tools

<http://socialmediadata.wikidot.com/>

Software for Web Scraping

<http://scraping.pro/software-for-web-scraping/>

Spinn3r - Indexing the Blogosphere

<http://docs.spinn3r.com/#overview>

SPSS Modeler

<http://developer.ibm.com/predictiveanalytics>

Squirro - Find, Remember, Organize and Share Important Information

<https://squirro.com/>

STACKS - Social Media Tracker, Analyzer, & Collector Toolkit at Syracuse

<https://github.com/bitlabsyr/stack>

TadaWeb - Clone and Amplify Human Intelligence for Web Data Collection and Analysis

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

Teracrawler – Cloud Based Web Crawling Software

<https://teracrawler.io/>

TextConverter 4

<https://www.simx.com/simx/TC-Overview.stp?>

TextRazor - Text Analysis Infrastructure

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

Topicgrazer - Graze On Web Pages and Documents

<http://www.topicscape.com/Topicgrazer/help.php>



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Trove - Privacy-Focused Bookmark Organizer  
<https://trovenow.com/>

UiPath – Web Data Extraction  
<https://www.uipath.com/guides/web-data-extraction>

Unit Miner - Web Data Extraction Software  
<http://www.unitminer.com/>

VietSpider  
<http://binhgiang.sourceforge.net/>

Visual Web Ripper – Data Extraction Software  
<http://www.VisualWebRipper.com/>

Visual Web Task  
<http://www.lencom.com/VisualWTSite.html>

Voogy – Anonymous Website Visitor Tracker  
<https://voogy.com/>

W3C Publishes Data Extraction Language (DEL) as W3C Note  
<http://xml.coverpages.org/ni2001-11-06-a.html>

Web Content Extractor  
<http://www.newprosoft.com/>

Web Data Extraction  
<https://www.wintask.com/web-data-extraction-feature/>

Web Data Extraction – Convert Websites Into Structured, Usable Data  
<https://www.import.io/product/extract/>

Web Data Extraction Services  
<https://webdataextractionservices.com/>

Web Data Extraction Software Data Toolbar  
<https://webdataextractionsoftwaredatatoolbar.en.softonic.com/>

Web Data Extractor  
<http://www.rafasoft.com/>



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Web Data Extractor

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

Web Data Extractor

<http://fivesmallq.github.io/web-data-extractor>

Web Data Extractor

<http://www.lantechsoft.com/web-data-extractor.html>

Web Data Extractors 2021

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

Web Data Guru – Web Data Extraction and Scraping Services

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

Web-Harvest – Open Source Web Data Extraction Tool

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

Webhose.io – Web Data For Your Business

<http://www.webhose.io/>

Web Robots – Web Scraping and Crawling

<https://webrobots.io/>

Web Scraper

<http://www.webscraper.io/>

Web Scraper Software Market Emerging Trends and Global Demands 2019 – 2024

<https://dailyresearchstack.com/2019/11/23/web-scraper-software-market-emerging-trends-and-global-demands-2019-to-2024/>

Web Scraping – Wikipedia

[https://en.wikipedia.org/wiki/Web\\_scraping](https://en.wikipedia.org/wiki/Web_scraping)

Website Data Extractor – Time to Rethink Web Scraping

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

Website Downloader

<https://websitedownloader.io/>



**November 2020 Zillman Column – Web Data Extractors 2021**

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

[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Website Extractor – Offline Browser  
<http://www.internet-soft.com/extractor.htm>

WebSunDew – Advanced Web Scraping Tool  
<http://www.websundew.com/>

Wikimedia Public Data Dumps  
[http://meta.wikimedia.org/wiki/Data\\_dumps](http://meta.wikimedia.org/wiki/Data_dumps)

WinAutomation  
<http://www.winautomation.com/>

XRay Web Scraping Tool  
<http://freecode.com/projects/xrayguibasedwebscrapingtool>

YaCy Web page Indexer  
<http://freecode.com/projects/yacy>



**November 2020 Zillman Column – Web Data Extractors 2021**

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

## **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/>

Business Intelligence Resources

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



November 2020 Zillman Column – Web Data Extractors 2021

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

[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

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/>

Games Resources

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

Genealogy Resources

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



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.



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/>

Intrapreneurial Resources

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

Journalism Resources

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

Knowledge Discovery

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



November 2020 Zillman Column – Web Data Extractors 2021

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

[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

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/>

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/>



November 2020 Zillman Column – Web Data Extractors 2021

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

[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

## World Wide Web Reference

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

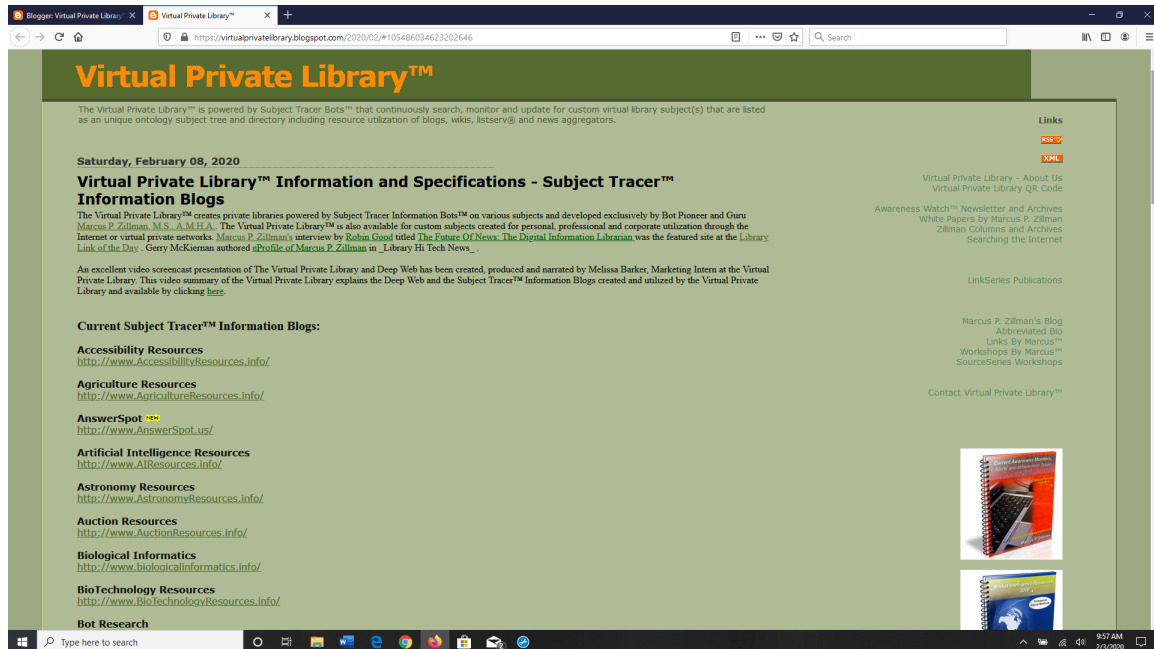


Figure 2: Virtual Private Library™

**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 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 (26,000+ Postings)

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

Marcus P. Zillman Abbreviated Bio

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



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

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/>

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

2020 Directory of Directories  
<http://www.2020DirectoryOfDirectories.com/>

2020 New Economy Resources  
<http://www.2020NewEconomy.com/>

2021 Guide to Finding Experts by Using the Internet  
<http://www.FindingExperts.info/>

2021 Guide to Finding People Resources and Sites  
<http://www.FindingPeople.info/>



**November 2020 Zillman Column – Web Data Extractors 2021**

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

2021 Guide to Internet Privacy Resources and Tools

<http://www.2021InternetPrivacy.com/>

2021 Guide to Searching the Internet

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

Academic and Scholar Search Engines and Sources 2020

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

Bots, Blogs and News Aggregators 2020

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

Business Intelligence Online Resources 2020

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

Cloud Computing Resources Primer 2020

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

Current Awareness Discovery Tools on the Internet 2020

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

Deep Web Research and Discovery Resources 2020 Online White Paper

<http://DeepWeb.us/>

eMarketing MiniGuide 2020

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

eReference Library Link Toolkit 2020

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

Finding Experts By Using the Internet 2021

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

Finding People Resources and Sites 2021

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

Healthcare Online Resources 2021

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



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

Knowledge Discovery Resources 2020

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

New Economy Resources 2020

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

Online Research Browsers and Data Visualization Tools 2020

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

Online Research Tools

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

Online Social Networking 2020

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

Searching the Internet 2021

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

Using the Internet As a Dynamic Resource Tool for Knowledge Discovery 2020

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

Web Data Extractors 2021

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

Web Guide for the New Economy 2020

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

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



November 2020 Zillman Column – Web Data Extractors 2021

<http://www.zillmancolumns.com/>  
[zillman@VirtualPrivateLibrary.com](mailto:zillman@VirtualPrivateLibrary.com)

eVoice: (800) 858-1462

© 2020 Marcus P. Zillman, M.S., A.M.H.A.

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.

### **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.

