

Web Data Extractors 2025

A White Paper Link Compilation

By

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

Extracting data from the World Wide Web (WWW) has become an important issue in the last few years as the number of web pages available on the visible Internet has grown to billions of pages with trillions of pages available from the invisible web. Tools and protocols to extract all this information have now come in demand as researchers as well as web browsers and surfers want to discover new knowledge at an ever increasing rate! As robots (bots) and intelligent agents are at the heart of many extraction tools I decided to create a compilation of the latest sources and sites that extract information from the web.

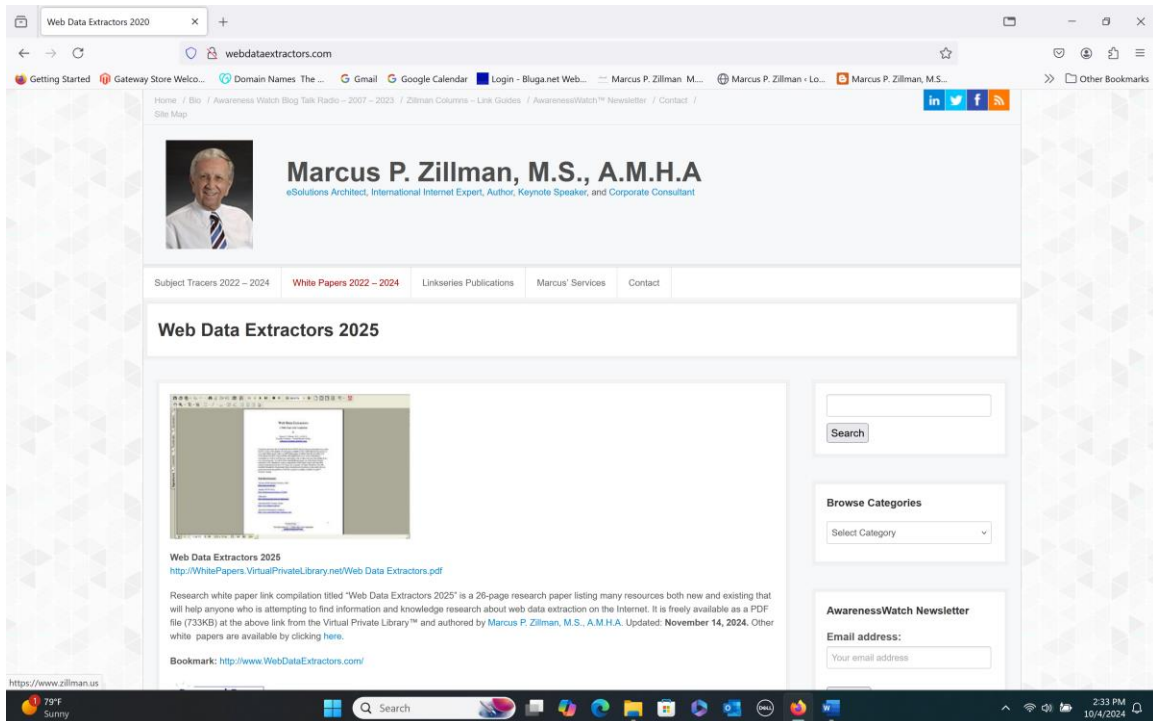


Figure 1: Web Data Extractors 2025



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Web Data Extractors 2025:

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

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

Agency – Robotic Process Automation (RPA) Software on Cloud for Data Scraping

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

Altair – Data Analytics and Artificial Intelligence (AI)

<https://www.altair.com/data-analytics/>

Anthracite

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

AnyBigData – Any Web Data You Want

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

Apify – Web Scraping Platform for Coders

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

ApiScrapy – AI-Driven Web Scraping & Data Labeling

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

Aristo - Answer Questions with a Knowledgeable Machine

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

Artificial Intelligence (AI) Discovery and Detection Tools 2024

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

artoo.js - The Client-Side Scraping Companion

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

AutoMate - Automate Data Extraction

<https://www.helpsystems.com/product-lines/automate>

Automated RSS Scraper Scripts

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

Automated Information Solutions

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

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/>

bli solutions Weather Predictive Analytics

<http://www.bli solutions.com/>

Bot Research 2023/2024

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

Browse.ai – Easiest Way to Extract and Monitor Data from Any Website

<https://www.browse.ai/>

BYU Data Extraction Research Group

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

Cogitum Co-Citer

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

Common Crawl

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

Crawl4AI (Async Version)

<https://github.com/unclecode/crawl4ai>

CrawlMonster

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

Crawly

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

Create a Crawler - Extract Data From an Entire Website

<https://www.import.io/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

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

Data Collection Infrastructure – Proxy Networks and Data Collection Tools
<https://brightdata.com/>

Data Excavator – Data Scraper for E-commerce
<https://data-excavator.com/>

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

DataHen – Empowering Enterprises with Clean Structured Web Data
<https://www.datahen.com/>

Data Mining Resources 2022/2023
<http://www.DataMiningResources.info/>

Data Miner – Powerful Web Scraping Tool for Professional Data Miners
<https://data-miner.io/>

Dataminr - Real-time AI Event and Risk Detection
<http://www.dataminr.com/>

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

Data Scraping Service – Get Public Data from the Web
<https://www.zyte.com/>

Data Scraping Services
<https://webdataextractionservices.com/>

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

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

Deep Web Research 2024
<http://www.DeepWebResearch.info/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

DEiXTo – Powerful Web Data Extraction Tool Based on W3C DOM

<http://deixto.com/>

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/>

DocuClipper – Data Extraction Software

<https://www.docuclipper.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 – Make Pricing Decisions with Confidence

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

File Information Tool Set (FITS)

<https://projects.iq.harvard.edu/fits>

FMiner – Web Scraping Software

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



Generative AI Resources 2024
<http://www.GenerativeAIResources.com/>

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/>

Hackaday – Tired of Web Scraping? Make the AI Do It
<https://hackaday.com/2023/04/09/tired-of-web-scraping-make-the-ai-do-it/>

Harkive – Data Collection – Multiple Sources/Single Database
<http://harkive.org/>

Helium Scraper – Extract Data from Any Website
<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>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Imagination Engines

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

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>

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/>

Knowledge Discovery Resources 2024

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Knowledge Graph Toolkit

<https://usc-isi-i2.github.io/kgtk/>

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/>

Mastodon Resources 2024

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

Metadata Extraction Tool

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

Minerazzi – Search and Mining Ecosystem

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

Mozenda – A dexi Brand - Comprehensive Web Data Gathering

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

<https://www.mozenda.com/mozenda-now-part-of-the-dexi-brand-family/>

Netlytic - Making Sense of Public Discourse Online

<https://netlytic.org/home/>

Newprosoft – Web Content Extractor

<http://newprosoft.com/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Octo – Expose Data From Any Database As Web Service

https://octoproject.github.io/octo-cli/?utm_campaign=Data_Elixir&utm_source=Data_Elixir_303

Octoparse – Easy Web Scraping for Anyone

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

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/>

Open MetaVerse Resources 2024

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

OpenSea – Web Data Extractor Pro

<https://www.OpenSea.io/>

Open Source Artificial Intelligence Agents (OSAIA) MiniGuide 2024

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

Open Source Intelligence (OSINT) Miniguide 2024

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

Outscrapr – Solutions for Accessing Public Information from the Internet for Lead Generation, Marketing, and Data Science

<https://Outscrapr.com/>

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

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

Page2API – The Ultimate Web Scraping API

<https://www.page2api.com/>

ParseHub – Free Web Scraper That Is Easy To Use

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Perplexity – Where Knowledge Begins

<https://www.Perplexity.ai/>

Priceonomics - Crawl Data From the Web

<http://priceonomics.com/>

Prompt Catalog 2024 for Artificial Intelligence (AI)

<http://www.PromptCatalog.ai/>

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/>

RAGFlow – Open-Source RAG (Retrieval-Augmented Generation) Engine Based on Deep Document Understanding

<https://ragflow.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/>

SalesHub – Find Your Ideal Prospects with Signals

<https://saleshub.ai/>

ScrapeForge

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

ScrapeGraphAI – LLM Powered Scraping

<https://colab.research.google.com/drive/1sEZBonBMGP44CtO6GQTwAIL0BGJXjtf?usp=sharing>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

ScrapeHero – Convert Websites Into Useful Data

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

Scraper

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

Scraper.ai – An AI Powered Web Scraper

<https://www.Scraper.ai/>

Scraper API – Proxy API for Web Scraping

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

Scraper: ChatGPT Plugin That Scrapes Websites with 1 Prompt

<https://artificialcorner.com/scraper-chatgpt-plugin-to-scrape-websites-with-1-prompt-56296e701edb>

ScrapeStorm – AI-Powered Web Scraping Tool and Web Data Extractor

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

ScrapeUp – Real Time Proxy API for Web Scraping

<https://scrapeup.com/>

ScrapingBot – API You Need for Efficient Scraping

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

ScrapingBytes – Performant Web Scraping API

<https://www.scrapingbytes.com>

ScrapingDog – Handles Millions of Proxies, Browsers and CAPTCHAs

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

ScrapingHub – Cloud Based Data Extraction Tool

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

Scraping Robot – Quality Web Scraping That You Can Count On

<https://scrapingrobot.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/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Screen-Scraper

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

Screen-Scraper – Web Data Extraction for Over Seventeen Years

<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>

Semantic Scholar - Free Scientific Literature Search and Discovery

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

Sensible Code

<http://sensiblecode.io/>

Sequentum – Unlock the World’s Largest Data Source

<https://sequentum.com/>

SerpApi – Google Search API

<https://serpapi.com/>

Sheet-Shaped Wikipedia: Turn Wikidata Into Spreadsheet-Ready Text Files

<https://lnkd.in/e2GwPY2y>

Simple Scraper – Extract Data From Any Website in Seconds

<https://simplescraper.io/>

Social Media Data Collection Tools

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

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/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

STACKS - Social Media Tracker, Analyzer, & Collector Toolkit at Syracuse
<https://github.com/bitslabsyr/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/>

TextSniper – Extract Text from Images and Other Digital Documents in Seconds
<https://textsniiper.app/>

Topicgrazer - Graze On Web Pages and Documents
<http://www.topicscape.com/Topicgrazer/help.php>

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/>

Vaazo – Web Bot That Can Scrape Data and Automate Tasks and More
<https://vaazo.com/>

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

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

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

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

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

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

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

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 Extractor and Scraper Tool
<https://www.webautomation.io/>

Web Data Extractors 2025
<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 – Tap Into Web Data Feeds at Scale
<http://www.webhose.io/>

Web Robots – Web Scraping and Crawling
<https://webrobots.io/>

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

WebScraping API – Leading REST API for Web Scraping
<https://www.webscrapingapi.com/>

Web Scraping – Wikipedia
https://en.wikipedia.org/wiki/Web_scraping

Web Scraping with Perl and ChatGPT
<https://proxiesapi.com/articles/web-scraping-with-perl-chatgpt>

Website Downloader
<https://websitedownloader.io/>

Website Extractor 10.52
<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

WinAutomation – Microsoft Power Automate
<http://www.winautomation.com/>

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

Xtract.io – Text Extration Mafde Easy
<https://www.xtract.io>

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

Zenscrape – An Elegant Web API for Ethical Data Extraction at Scale
<https://zenscrape.com/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 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.co/>

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/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 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/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 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/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 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/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

World Wide Web Reference
<http://www.WWWReference.info/>

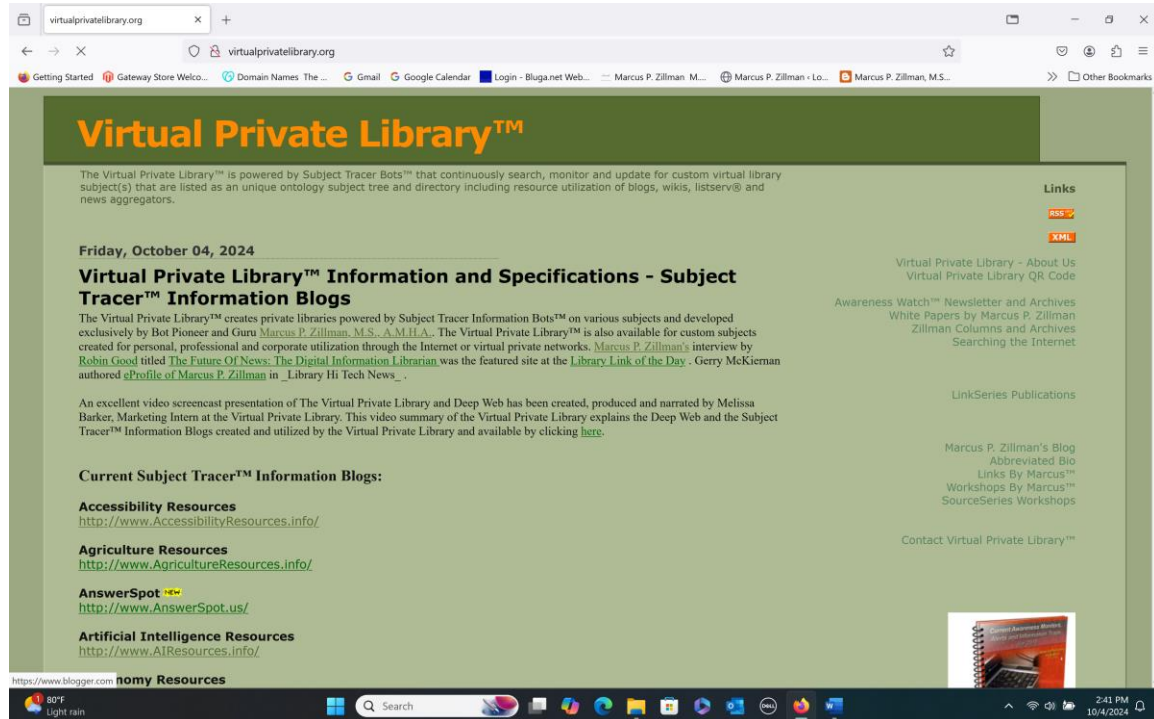


Figure 2: Virtual Private Library™

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 (30,000+ Postings)

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Marcus P. Zillman Abbreviated Bio
<http://www.zillman.info/>

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.:

2022/2023 Guide to Finding Experts by Using the Internet
<http://www.FindingExperts.info/>

2022/2023 Guide to Finding People Resources and Sites
<http://www.FindingPeople.info/>

2022/2023 Guide to Internet Privacy Resources and Tools
<http://www.2022InternetPrivacy.com/>

2024 Directory of Directories
<http://www.2024DirectoryOfDirectories.com/>

2024 Green Files
<http://www.GreenFiles.info/>

2024 Guide to Searching the Internet
<http://www.SearchingTheInternet.info/>



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

2024 New Economy Resources

<http://www.2024NewEconomy.com/>

2024 Publications by Marcus P. Zillman, M.S., A.M.H.A.

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

2024 Reference Resources

<http://www.2024ReferenceResources.com/>

Academic and Scholar Search Engines and Sources 2024

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

Artificial Intelligence (AI) Discovery and Detection Tools 2024

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

Bots, Blogs and News Aggregators 2024

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

Business Intelligence Online Resources 2022/2023/2024

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

Cloud Computing Resources Primer 2025

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

Current Awareness Tools 2025

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

Deep Web Research and Discovery Resources 2024 Online White Paper

<http://DeepWeb.us/>

eMarketing MiniGuide 2024

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

eReference Library Link Toolkit 2022/2023

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

Fact Check Resources Miniguide 2024

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

Finding Experts By Using the Internet 2022/2023

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Finding People Resources and Sites 2022/2023

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

Generative AI Resources 2024

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

Healthcare Bots and Subject Directories 2024

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

Healthcare Online Resources 2022/2023

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

Knowledge Discovery Resources 2024

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

New Economy Resources 2025

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

New Normal StartUp Resources 2024

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

Online Research Browsers and Data Visualization Tools 2024

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

Online Research Tools 2023/2024

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

Online Social Networking 2022/2023

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

Open DataSets 2024

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

Open Educational Resources (OER) Sources 2024

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

Open MetaVerse Resources 2024

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

Open Source Artificial Intelligence Agents (OSAIA) MiniGuide 2024

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



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

Open Source Intelligence (OSINT) Miniguide 2024
<http://www.OSINTminiguide.com/>

Prompt Catalog 2024 for Artificial Intelligence (AI)
<http://www.PromptCatalog.ai/>

Searching the Internet 2024
<http://www.SearchingTheInternet.info/>

Social Informatics 2022/2023/2024
<http://www.SocialInformatics.net/>

Subject Tracers 2022/2023/2024
<http://www.SubjectTracers.com/>

Using the Internet As a Dynamic Resource Tool for Knowledge Discovery 2024
<http://www.zillman.us/white-papers/using-the-internet-as-a-dynamic-resource-tool-for-knowledge-discovery/>

Web Data Extractors 2025
<http://www.WebDataExtractors.com/>

Web Guide for the New Economy 2024
<http://www.WebGuideNewEconomy.com/>

White Papers 2022/2023/2024 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.



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

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.



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.

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.



Web Data Extractors 2025 – A White Paper Link Compilation

[Updated: November 14, 2024]

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

zillman@VirtualPrivateLibrary.com

239-206-3450

© 2007 – 2024 Marcus P. Zillman, M.S., A.M.H.A.