**Mayuri Shridatta Yerande**
**Class: D15B**
**Roll no: 70**

# EXPERIMENT - 1

**Aim**: Study Any 1 tool in each
1. Study web analytics using open source tools like Matomo, Open Web Analytics, AWStats, Countly, Plausible.
2. Study Semantic Web Open Source Tools like Apache TinkerPop, RDFLib, Apache Jena, Protégé, Sesame.

## Theory:

**What is web analytics?**
Web analytics is the process of analyzing the behavior of visitors to a website. This involves tracking, reviewing and reporting data to measure web activity, including the use of a website and its components, such as web pages, images and videos.

**Analytics can help in the following ways:**

- Determine the likelihood that a given customer will repurchase a product after purchasing it in the past.
- Personalize the site to customers who visit it repeatedly.
- Monitor the amount of money individual customers or specific groups of customers spend.
- Observe the geographic regions from which the most and the least customers visit the site and purchase specific products.
- Predict which products customers are most and least likely to buy in the future.

**Process of web analytics:**

- Setting goals.
- Collecting data.
- Processing data.
- Identifying key performance indicators (KPIs)
- Developing a strategy.
- Experimenting and testing.

**What is semantic web?**

The Semantic Web is a mesh of data that are associated in such a way that they can easily be processed by machines instead of human operators. It can be conceived as an extended version of the existing World Wide Web, and it represents an effective means of data representation in the form of a globally linked database.

The Semantic Web can be considered a vision for the future in which data could be quickly interpreted by machines, allowing them to carry out numerous tedious tasks related to discovering, blending, and taking action on the information available on the Web.

By supporting the inclusion of semantic content in Web pages, the Semantic Web targets the conversion of the presently available Web of unstructured documents to a Web of information/data.

The key goal of the Semantic Web is to trigger the evolution of the existing Web to enable users to search, discover, share and join information with less effort. Humans can use the Web to execute multiple tasks, such as booking online tickets, searching for different information, using online dictionaries, etc.

Even so, machines are not able to carry out any of these tasks without human intervention because Web pages are made to be read by humans, not machines. The Semantic Web can be considered a vision for the future in which data could be quickly interpreted by machines, allowing them to carry out numerous tedious tasks related to discovering, blending, and taking action on the information available on the Web.

| Web 1.0 | Web 2.0 | Web 3.0 |
|---|---|---|
| Mostly Read-Only | Wildly Read-Write | Portable & Personal |
| Company Focus | Community Focus | Individual Focus |
| Home Pages | Blogs / Wikis | Lifestreams / Waves |
| Owning Content | Sharing Content | Consolidating Content |
| Web Forms | Web Applications | Smart Applications |
| Directories | Tagging | User Behavior |
| Page Views | Cost Per Click | User Engagement |
| Banner Advertising | Interactive Advertising | Behavioral Advertising |
| Britannica Online | Wikipedia | The Semantic Web |
| HTML / Portals | XML / RSS | RDF / RDFS / OWL |

**Tools in web analytics:-**

- Matomo
- Open Web Analytics
- AWStats
- Countly
- Plausible

**Tools in semantic web:-**

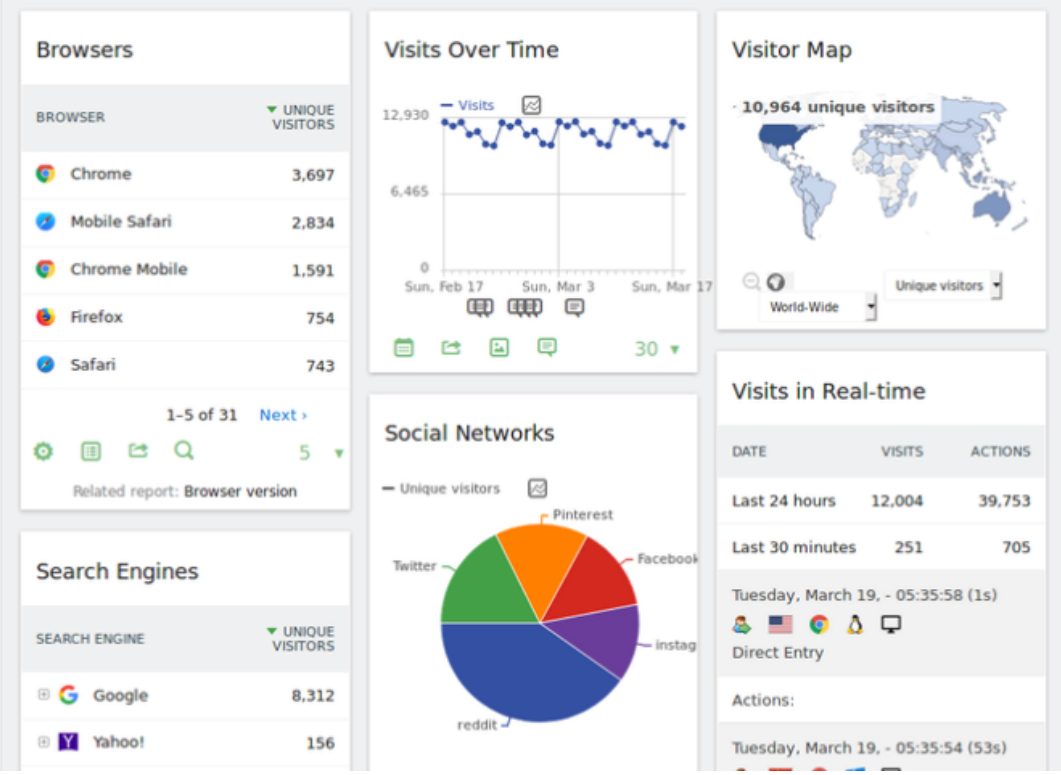
- Apache TinkerPop
- RDFLib
- Apache Jena
- Protégé
- Sesame.

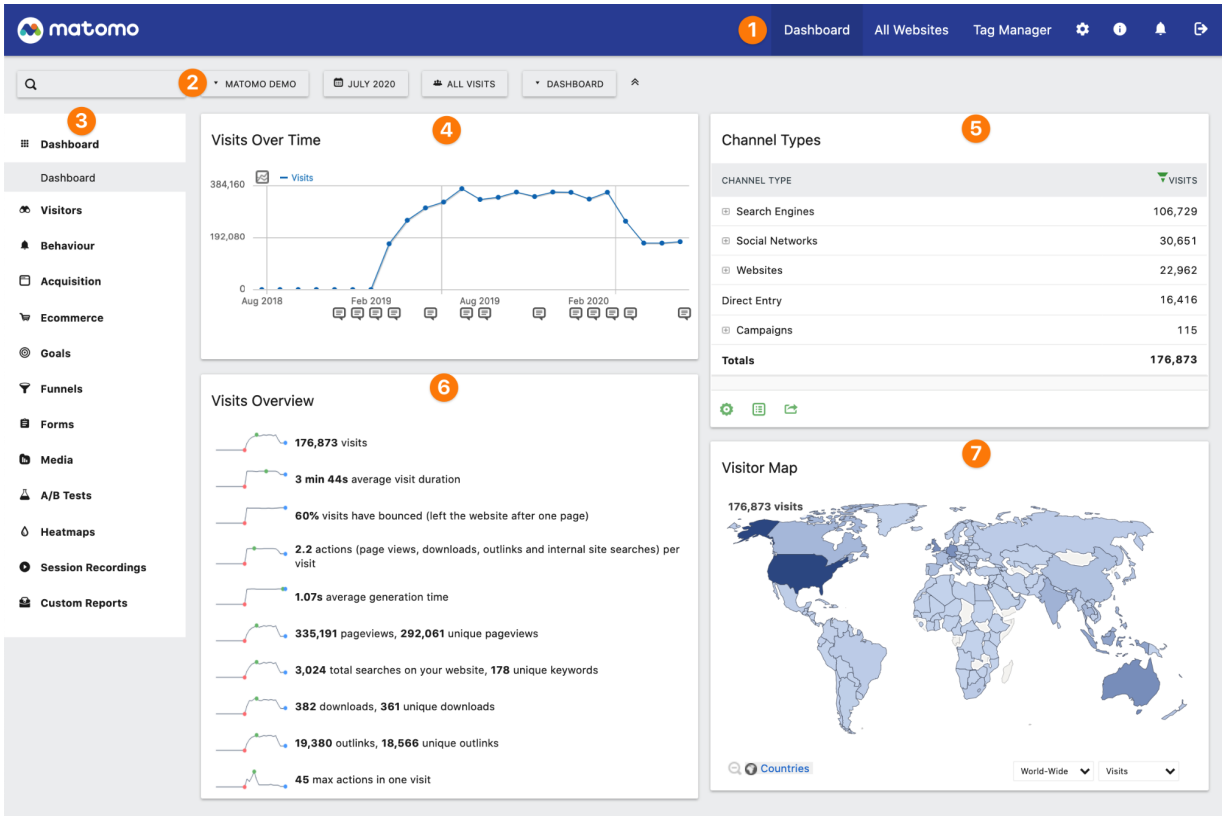
## Web Analytics Tool:- Matomo

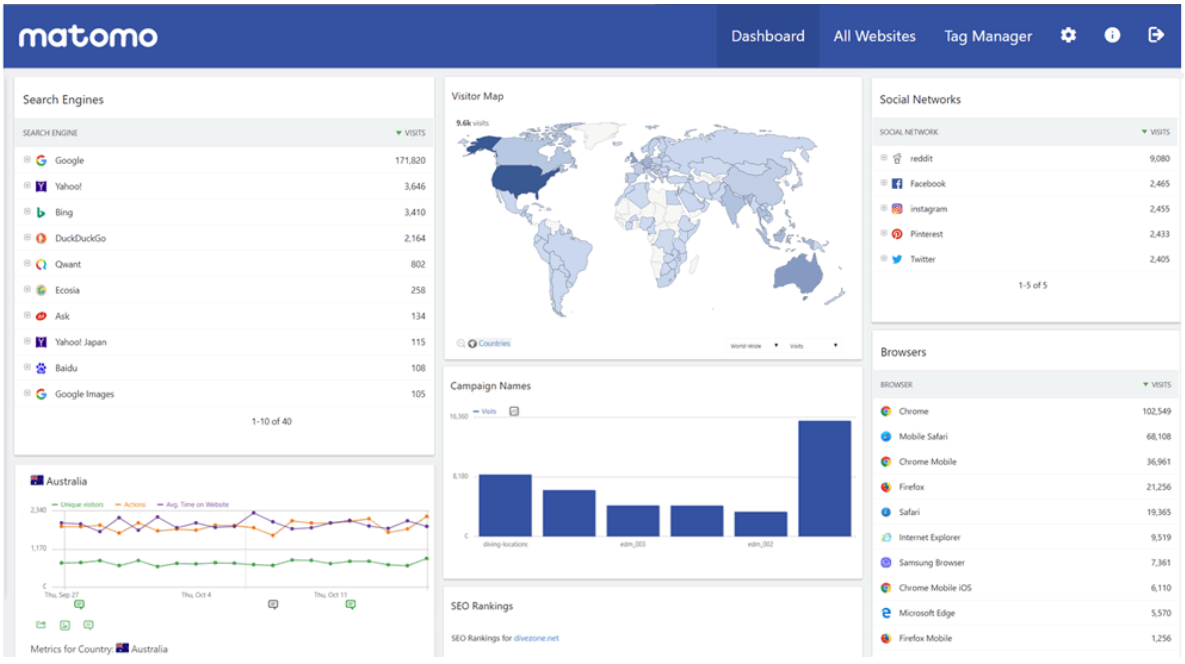
Matomo is a downloadable, Free (GPL licensed) web analytics software platform. It provides detailed reports on your website and its visitors, including the search engines and keywords they used, the language they speak, which pages they like, the files they download and so much more.

What data does Matomo track?

- User IP address (see also: IP anonymisation)
- Optional User ID.
- Date and time of the request.
- Title of the page being viewed (Page Title)
- URL of the page being viewed (Page URL)
- URL of the page that was viewed prior to the current page (Referrer URL)
- Screen resolution being used.
- Number of Visits
- Number of users
- Social media engagement tracking



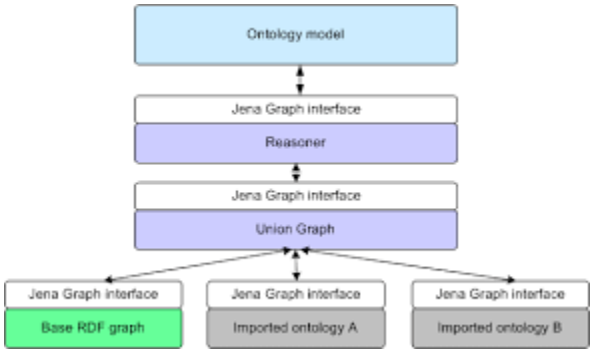

matomo
1
Dashboard
All Websites
Tag Manager
2
MATOMO DEMO
JULY 2020
ALL VISITS
DASHBOARD
3
Dashboard
Dashboard
Visitors
Behaviour
Acquisition
Ecommerce
Goals
Funnels
Forms
Media
A/B Tests
Heatmaps
Session Recordings
Custom Reports
Visits Over Time
4
384,160
192,080
0
Visits
Aug 2018
Feb 2019
Aug 2019
Feb 2020
Channel Types
5
CHANNEL TYPE
VISITS
Search Engines 106,729
Social Networks 30,651
Websites 22,962
Direct Entry 16,416
Campaigns 115
Totals 176,873
Visits Overview
6
176,873 visits
3 min 44s average visit duration
60% visits have bounced (left the website after one page)
2.2 actions (page views, downloads, outlinks and internal site searches) per visit
1.07s average generation time
335,191 pageviews, 292,061 unique pageviews
3,024 total searches on your website, 178 unique keywords
382 downloads, 361 unique downloads
19,380 outlinks, 18,566 unique outlinks
45 max actions in one visit
Visitor Map
7
176,873 visits
Countries
World-Wide
Visits


matomo
Dashboard
All Websites
Tag Manager
Search Engines
SEARCH ENGINE
VISITS
Google 171,820
Yahoo! 3,646
Bing 3,410
DuckDuckGo 2,164
Qwant 802
Ecosia 258
Ask 134
Yahoo! Japan 115
Baidu 108
Google Images 105
1-10 of 40
Visitor Map
9.6k visits
Countries
Social Networks
SOCIAL NETWORK
VISITS
reddit 9,080
Facebook 2,465
instagram 2,455
Pinterest 2,433
Twitter 2,405
1-5 of 5
Campaign Names
Visits
Browsers
BROWSER
VISITS
Chrome 102,549
Mobile Safari 68,108
Chrome Mobile 36,961
Firefox 21,256
Safari 19,365
Internet Explorer 9,519
Samsung Browser 7,361
Chrome Mobile iOS 6,110
Microsoft Edge 5,570
Firefox Mobile 1,256
Australia
Metrics for Country: Australia
SEO Rankings
SEO Rankings for divezone.net

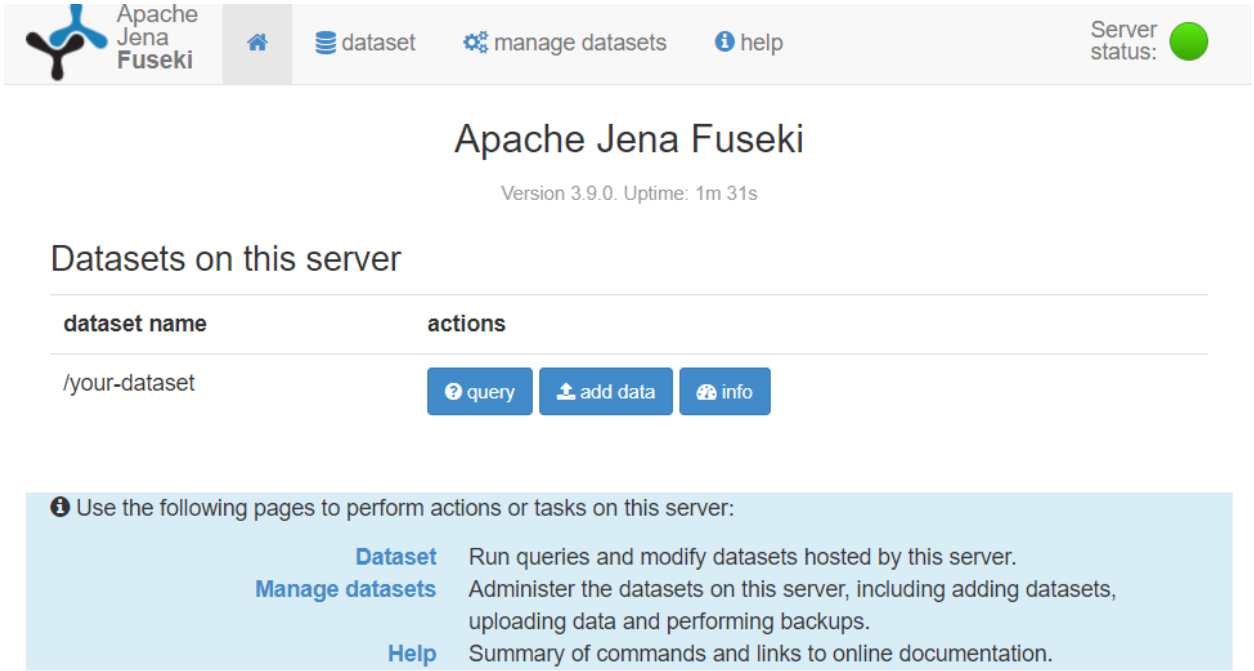
## Semantic Web tool:- Apache Jenna

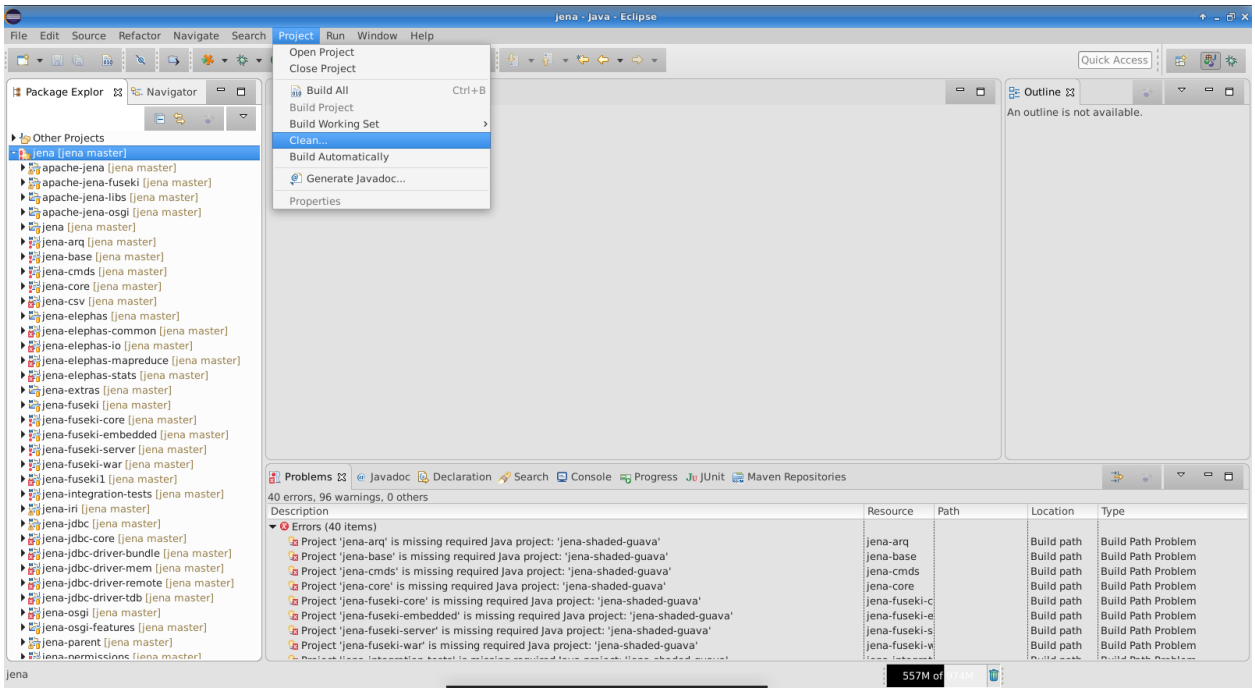
Jena is a Java framework for building Semantic Web applications. It provides an extensive Java library for helping developers develop code that handles RDF, RDFS, RDFa, OWL and SPARQL in line with published W3C recommendations.

Jena was originally developed by researchers in HP Labs, starting in Bristol, UK, in 2000. Jena has always been an open-source project, and has been extensively used in a wide variety of semantic web applications and demonstrators.




In 2009, HP decided to refocus development activity away from direct support of development of Jena, though remaining supportive of the project's aims.
Jena includes a rule-based inference engine to perform reasoning based on OWL and RDFS ontologies, and a variety of storage strategies to store RDF *triples* in memory or on disk.




**CONCLUSION:** We successfully studied about the Semantic web and web analytics. We studied the matomo and apache Jenna where we saw what exactly do they do, and also implemented them.