

Welcome to ezPAARSE 'em

Towards an increasingly refined understanding of activity

My name is Derek Zoladz

Library systems analyst at OhioNET

And you can find these slides at the link below

Working with EZproxy logs can be quite challenging

Extracting meaningful data is not a simple task

Particularly, extracting data that is insightful and actionable can be burdensome

So, why is that?

We're dealing with large amounts of data. EZproxy logs can easily contain 10s of thousands of entries

With all of those entries, determining the signal from the noise can be challenging

Vendors use differing URL structure, so building and maintaining custom scripts to extract useful data can become, well...

Quite overwhelming

Enter ezPAARSE. ezPAARSE addresses several of these challenges

It provides a community-maintained set of parsing tools for over 250 major publishing platforms

It automatically filters and removes unimportant lines from your log files

It uses data in your EZproxy logs to enhance and build new information

And since ezPAARSE produces a uniform output file, you can supplement and cross-reference the results against other institutional data, to obtain an even deeper level of understanding and insight.

Let's see it in action

<https://ezparse.opal-libraries.org>

[DEMONSTRATION]

The first step in the process is to create a new account

Once you have an account, go ahead and sign in

Before we get started, I want to point out a specific section in the documentation.

It's crucial to have an understanding of Consultation and Access Events

This chart defines the properties in an ezPAARSE output file

I would also recommend reviewing the Resource Types (or, rtypes) to grasp the categories of things ezPAARSE understands

Processing your EZproxy log files is straightforward

Drag and drop your log files into ezPAARSE and click continue

In the predefined configuration dropdown, select OhioNET from the menu of option

Choosing OhioNET will:

In the input section,

- Set the log type to EZproxy
- Populate the log format expression to match OhioNET's log format.

In the output section,

- Ensure that the output file is TSV, or tab separated values
- Enable the COUNTER-compliant JR1 report to be generated

To begin processing the log files, click 'start processing'

ezPAARSE will begin to read each line of the input file, extract any useful data points, and begin to generate the parsed output file.

Once complete, click 'download results' and open the file in Microsoft Excel, or Apple's Numbers, to view the results

You can clearly see:

- Date and Time of the event
- Platform that was Accessed
- Media Type that was consumed
- Latitude and Longitude of the Originating Request
- The identifier used to login to the EZproxy session
- And the browser being used

and when available, data enrichment will occur. You'll be able see the specific titles being accessed.

With this data, you can see:

- Trends in remote usage, throughout the semester or by time-of-day
- Identify the types of media that are most consumed on a subscription platform
- Target the development of the library's web properties around the browsers that your users actually prefer

But maybe more importantly, it's possible to:

- Correlate logins and usage with specific campus departments or programs

Using supplemental institutional data

If you have any questions, or if OhioNET hosts your instance of EZproxy and you want to inquire about gaining access to your log files

Send us an email at support@ohionet.org

Otherwise

Happy Parsing!