LibStats: There and Back Again

Presented By: Elizabeth Peele Mumpower, Systems Librarian
             Alex Cooper, Lead Applications Analyst
             Lisa Hamlett, Applications Analyst II
Road Map

● Introduction

● Part I:
  ○ Why this project?
  ○ How did we start it?

● Part II:
  ○ Engaging With the ezPaarse Open Source Community
  ○ Making ezPaarse Your Own

● Part III:
  ○ Using Tableau to compile and present your unique data
  ○ Some examples of our data
Part 1.
Beginning the journey

Origin, project management, and architecture
The Origins of LibStats

Have

- COUNTER
- Unparsed proxied logs
- Proxy both on & off campus

Need

- Data from resources without COUNTER
- Easier way to visualize stats
- Way to match access with demographic data for a more well-rounded view
ezPAARSE
Project Management

● Multiple moving components
  ○ Setting up ezPaarse
  ○ Architecting data loads
  ○ Designing & Implementing database
  ○ Setting up Tableau
  ○ Ensuring communication with stakeholders
  ○ Prioritization
  ○ etc

● Use GitHub to track everything
  ○ Private repo for initial setup and changes/tweaks to architecture
  ○ Public repo forked from ezpaarse-platforms
Private Repo
Architecture

- **Extract:**
  - Different data sources

- **Transform:**
  - Filter
  - Match
  - Anonymization
  - Date transformation

- **Load:**
  - Oracle Database
Part 2.
On the Road
Engaging with the ezPaarse community and making ezPaarse your own
Tracking the State of ezPaarse

- Couperin and the project members
- Trello: ezPaarse / AnalogIST
- The Couperin project AnalogIST webpage
  - http://ang.couperin.org/
Contributing to ezPaarse Part 1

- Analyze your EZProxy logs
- Create a Trello card for the platform
- Add links from the EZProxy logs to AnalogIST
Contributing to ezPaarse Part 2

- Creating your own parsers
- Use platform-init to create your platform’s parser
- Modify the parser.js template to suit your needs
- Build out your test.csv
- Commit your code to ezPaarse on GitHub
  - [https://github.com/ezpaarse-project/ezpaarse-platforms](https://github.com/ezpaarse-project/ezpaarse-platforms)
- Since 01/2018 we have added 53 parsers locally and 35 globally
  - NOTE: We do not reparse old logs when we add a new parser
Tools

- regex
  - Tooltip: [https://regex101.com/](https://regex101.com/)
- JavaScript
- ezPaarse community
Part 3.
Telling Your Story

Using Tableau to compile and present your unique data
Preparing the Data Connection

- Tableau Public vs. Tableau Desktop
  - Same functionality, but access to data in Tableau Public cannot be restricted and Tableau owns the data
- Connect Tableau Desktop to Tableau Server
- Connect Tableau Desktop to Database with Tableau Connector
- Publish data source to Tableau Server
- Schedule a regular extract (not a live connection) of the data source
- Create a workbook connected to data source
  - Pro tip: You can connect multiple workbooks to a single data source from the server
- Publish workbook to Tableau Server
LibStats vs. COUNTER

- LibStats is *not* a replacement for COUNTER reports
  - COUNTER data describe *what* resources are being accessed
  - LibStats data describe *who* is accessing the resources

- LibStats data for resources is inconsistent
  - Every website handles links and resources uniquely
  - Parser writers have different ideas of what constitutes a certain type of resource

- Both LibStats and COUNTER reports are needed to get a full picture of collection usage
LibStats Data

- LibStats looks at these specific data:
  - Actions, Data (Bytes), Unique Sessions, Unique Title IDs, Unique Unit IDs, & Unique User IPs
- Some data is incomplete because many parsers only show initial access
- The most useful data defined
  - Actions: Anything a user does (click, open a file, view a TOC, etc.)
  - Unique Sessions: Number of unique logins into ezProxy
  - Unique User IPs: Number of unique IP addresses that have been logged into ezProxy
That’s nice and all…
But let’s see the data!
Actions by Month by Person Type

Click a person type in the legend to highlight that type only.

Select Measure
- Actions
- Total Data (Bytes)
- Unique Sessions
- Unique Title IDs
- Unique Unit IDs
- Unique User IPs

Person Type (group)
- Student & Student St.
- Faculty
- Staff & Staff Student
- Unknown
- Walk-In Stations
- Retired
- Sponsored
- Pre-start
- Student Applicant
- Administrative

Recorded Date
10/1/2016 12:00:00 AM
and Null values
Questions?
THANK YOU!

CONTACT:

Elizabeth Peele Mumpower, Systems Librarian

epeele@emory.edu

Alex Cooper, Lead Applications Analyst

acoope5@emory.edu

Lisa Hamlett, Applications Analyst II

lhamlet@emory.edu
Credits & Links

- Presentation template by [SlidesCarnival](https://trello.com/b/wEaLnz8d/ezpaarse-analogist-suivi-des-plateformes)
- [https://github.com/ezpaarse-project/ezpaarse-platforms](https://github.com/ezpaarse-project/ezpaarse-platforms)
- [https://regex101.com/](https://regex101.com/)
- [https://github.com/emory-libraries/ezpaarse-platforms](https://github.com/emory-libraries/ezpaarse-platforms)