



# Gold Rush Analytics

**George Machovec**, Executive Director  
Colorado Alliance of Research Libraries

October 6, 2021

[george@coalliance.org](mailto:george@coalliance.org)

# Colorado Alliance of Research Libraries

- Incorporated in 1981 at 501c3 non-profit
- 16 member libraries (15 academic, 1 public) – one out of state library (University of Wyoming)
- Programmatic Focus – all are opt-in
  - E-resource licensing – 250 contracts (>\$19 million)
  - Prospector union catalog – 100+ libraries, 15+ million unique MARC records
  - Gold Rush
  - Alliance Shared Print Trust (launched 2015)
- A history of innovation and software development



# Gold Rush Analytics Technical Overview

- Developed and hosted by the Colorado Alliance of Research Libraries
- Highly modified Blacklight/Solr implementation
  - Uses mostly open-source software with a few exceptions (e.g. charting)
- Servers located at the University of Denver
  - Much less expensive than AWS hosting at this point
- Code is backed-up in GitHub
- Offered as a service by the Colorado Alliance at cost
  - No startup fees
  - Friendly contract options



# Why We Created the System

- Alliance Shared Print Trust
  - Covers academic libraries in Colorado and Wyoming
  - Broad agreement (MOU) signed by participating libraries
  - Circulating monographs policy, Serials policy, Disclosure policy
  - 25-year commitment, reviewed every 5 years
- An analysis tool was needed that was affordable and flexible
  - Real time comparative analysis at-scale
  - Can be regularly updated, as needed, by participating libraries
  - Creates a regional database of retention commitments
- Commercial tools are excellent but were too expensive for member libraries

# Gold Rush Analytics – 2 modules

## Two comparison modules in the system

- Journal comparison
  - Originally developed in 2003 to compare aggregators, publisher lists, indexing/abstracting services, local holdings (over 1,800 title lists)
  - Comparisons on ISSN/eISSN
  - Libraries may ingest holdings from an ILS, ERMS, discovery layer, RapidILL or other sources
  - Metadata is K-BART or thin metadata from vendors
- MARC record comparison tool
  - Full MARC record ingest
  - With special match key

# Gold Rush Design Principles

- Must be scalable to work from the smallest to largest libraries
- Must work in real-time
- Must support single library and comparative analytics
- Must retain MARC records from each site
- Must support export of MARC, XML and delimited (Excel)
- Must show what's unique and overlapping, including visualizations
- Must be able to have searches or facets to work with virtually any fields in a MARC record
- Matching algorithm must be made from elements in a MARC record but **NOT** depend on OCLC #, ISBN, ISSN
- Must be easy-to-use!

# Loading and updating

- Can load MARC21 Records from any system
  - Any item information should be in a 9XX field (not separate LHR records at this time)
  - Item-level retention commitments in MARC 583
- Suggested that full database updates be done regularly if needed
  - Deletes are currently handled through a complete database refresh
- SFTP (secure FTP) is used to deposit records in either a “full” or “updates” directory
- Profiling spreadsheet needed if branch/locations are loaded from item (holdings) records

# Selected Elements in Match Key

- **Title**
  - 245 \$a \$b (first 70 characters)
- **General Media Description**
  - 245 \$h
- **Type**
  - ' \_ ' Leader
- **Title Part**
  - 245 \$p
- **Title Number**
  - 245 \$n
- **Publication Year**
  - 008 or 260 or 264 \$c
- **Pagination**
  - 300 \$a
- **Edition Statement**
  - 250 \$a
- **Publisher Name**
  - 260 or 264 \$b
- **NO matching is done with ISBN, ISSN or OCLC #** (although searchable)
- **Match key ends with an “e” or “p” to separate electronic and print**



Save search Load saved search Export records 56 found . Displaying records 1 to 25 Previous Next Pivot reports

Limit your search... Apply Show 25 entries, sorted by Relevance in descending order. Apply Sort

- Format
- Publication year
- Subject heading
- Language
- LC call number
- Dewey call number
- Library
- Overlap
- Branch Code
- Location Code

**Title:** Field and service robotics  
**Author:** International Conference on Field and Service Robotics  
**Format:** Electronic, Book  
**Subject heading:** Mathematics, Engineering, Systems theory  
**Language:** English  
**Region:**  
**Publication Date:** 2010  
**Publisher:** Springer  
**OCLC #:**  
**Library:** [University of Wyoming](#), [University of Denver](#), [Auraria](#), [denver--all](#),  
**Branch:**  
**Circ. Count #:**  
**Match Key:** [field\\_and\\_service\\_roboticsresults\\_of\\_the\\_7th\\_international\\_conference\\_elect2010\\_\\_\\_\\_\\_spri...](#)

**Title:** Field and service robotics  
**Author:** International Conference on Field and Service Robotics  
**Format:** Electronic, Book  
**Subject heading:** Robotics, Robots, Industrial, Electronic books  
**Language:** English  
**Region:**  
**Publication Date:** 2008  
**Publisher:** Springer  
**LC #:** TJ210.3 .F545 2008  
**OCLC #:**  
**Library:** [Colorado School of Mines](#), [CSU Fort Collins](#),  
**Branch:**  
**Circ. Count #:** Westernuca--all--  
**Match Key:** [field\\_and\\_service\\_roboticsresults\\_of\\_the\\_6th\\_international\\_conference\\_\\_\\_\\_\\_2008\\_\\_\\_\\_200spri...](#)

**Title:** Field and Service Robotics  
**Author:** Zelinsky, Alexander

Three libraries have matched up. Select the library to see their MARC record (all are retained)

A master indexing record is selected based on encoding level and order of processing. MARC records from all sites are linked to it

# Complete Marc Record

### MatchKey:

"field\_and\_service\_roboticsresults\_of\_the\_7th\_international\_conference\_elect2010\_\_\_\_sprina\_\_\_\_international\_confer\_\_\_\_\_e"

Match Key

**ID:** S\_EX\_WY\_UWY\_NO\_ID\_1912600

**Title:** Field and service robotics

**Author:** International Conference on Field and Service Robotics (7th : 2009 : Cambridge, MA, USA)

**Format:** Electronic, Book

**Publication Date:** 2010

**Item Circ Count:**

**Item Branch:**

**Item Location:** ,

**Library :** University of Wyoming;

**LC code:** TJ210.3 .I58 2009

**LC sort:** 6073667050985750531

View MARC

000 01555nam a2200361uu 4500

003 NhCcYBP

005 20100819143826.1

006 m d

007 cr n

008 100819s2010 gw s 000 0 eng d

020 |a9783642134081 (electronic bk.)

040 |aNhCcYBP|cNhCcYBP

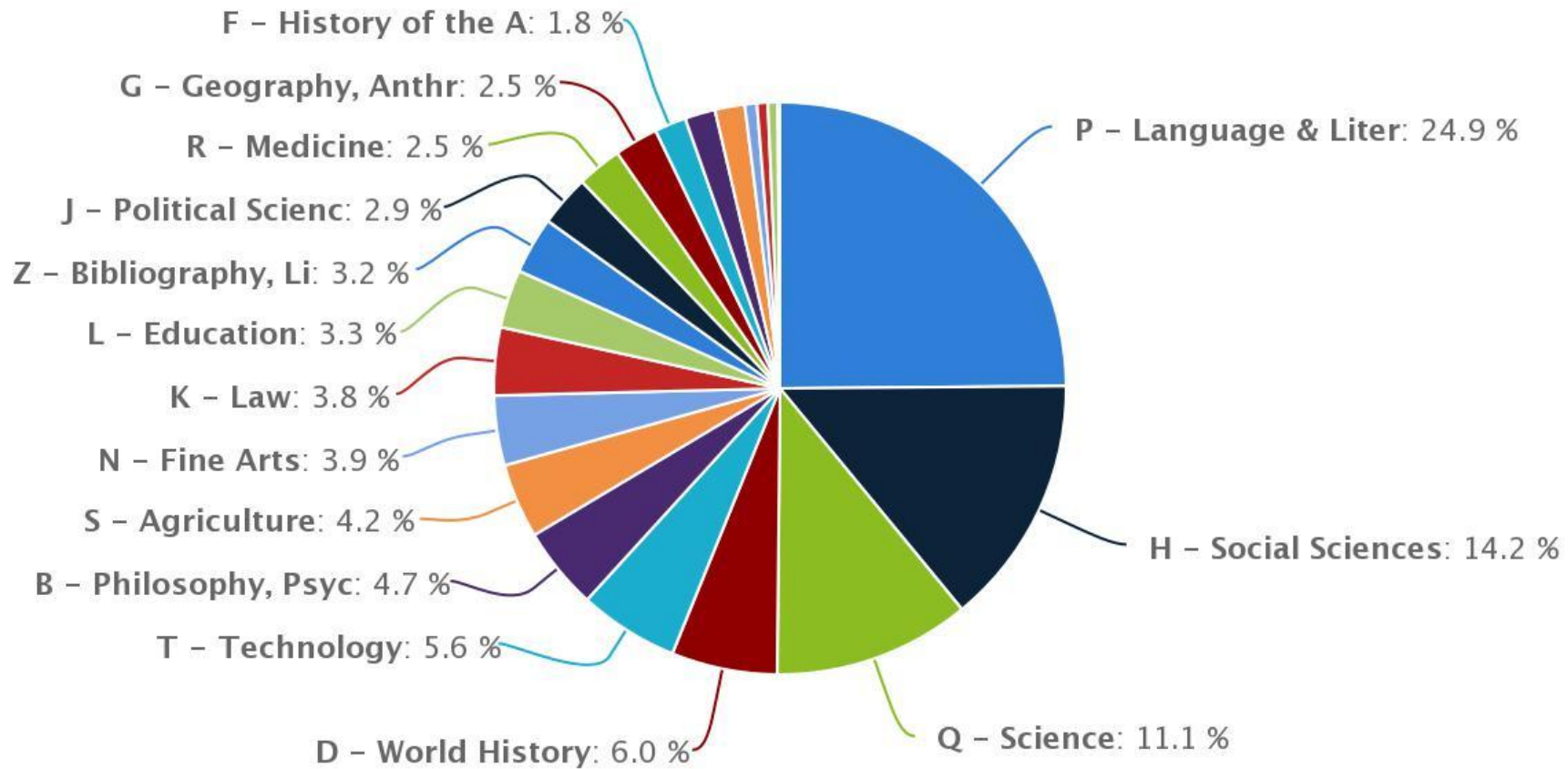
050 4|aTJ210.3|b.I58 2009

000 01555nam a2200361uu 4500

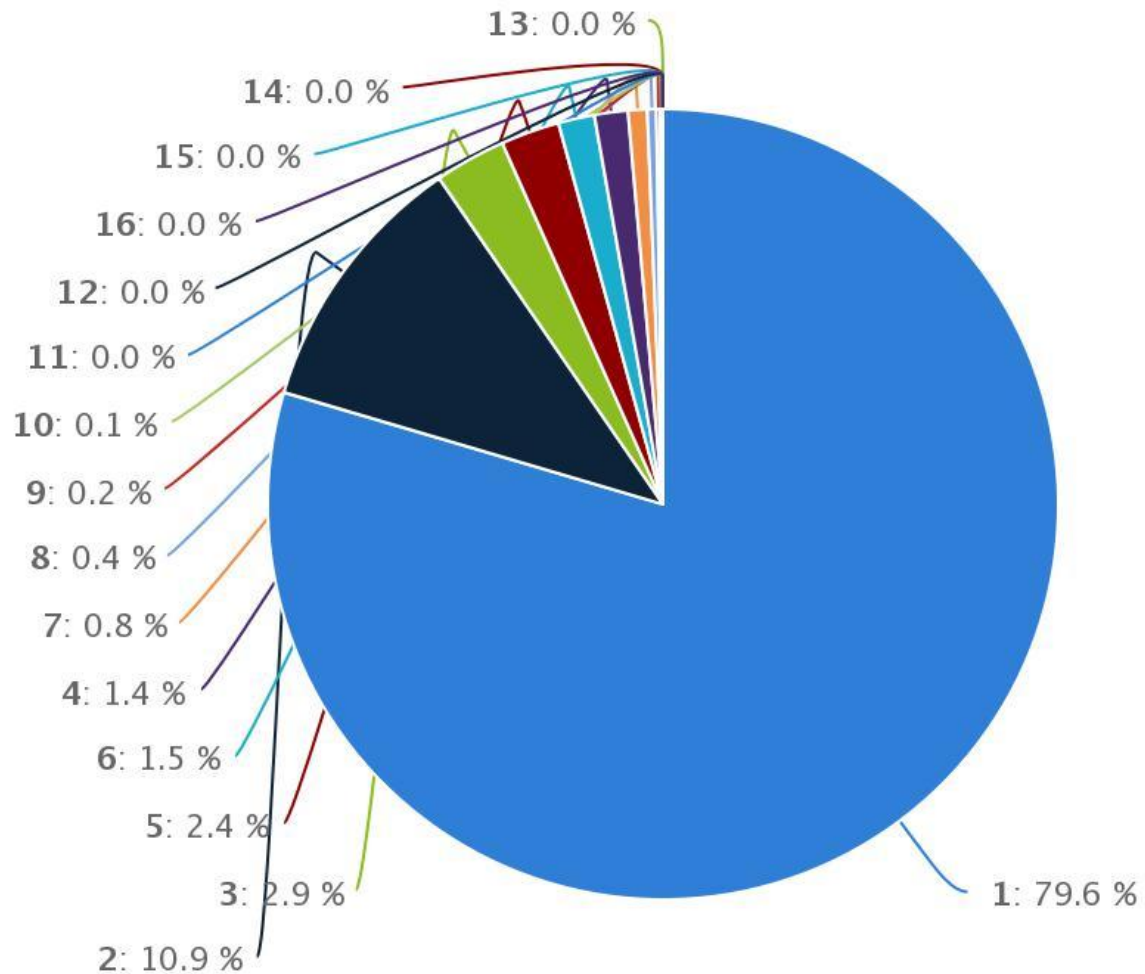
# Sample Visualization

## Colorado Alliance Retention Commitments by LC Call Number

LC call number facet



## Overlap facet



**Uniqueness (Overlap)  
of the global EAST  
retention  
commitments (n=6.8  
million)**

These may be  
fine-tuned by any of  
the available facets or  
searches

# Sample Use Cases for GR Analytics



- A librarians view to your cataloging records
- Shared print programs
  - What to weed or add to shared print program
  - Analyzing overlap of retention commitments
- Collective Collection analytics. An interactive data lake
- Determining publisher backfile holdings for licensing negotiations
- Space analysis for a remodel or new building program
- Adding a new program on campus and want to compare holdings with an established institution with the same program
- Exports for cooperative programs, Google book scanning, aggregating records for discovery layers, etc.

# Links of Interest

- Article in Collaborative Librarianship (April 2016) -  
<http://digitalcommons.du.edu/collaborativelibrarianship/vol8/iss1/7/>
- FAQ about Gold Rush -  
<https://www.coalliance.org/faq-library-content-comparison-system>

# Who to Contact

- Gold Rush licensing
  - George Machovec, [George@coalliance.org](mailto:George@coalliance.org)
  - Rose Nelson, [rose@coalliance.org](mailto:rose@coalliance.org)
- For general questions and Gold Rush support:  
[goldrush@coalliance.org](mailto:goldrush@coalliance.org)