AGUA Development Phase 7

Western Regional Storage Trust

Final Report

Summary

AGUA is a home-grown tool, maintained by the AGUA Development Team at the California Digital Library, and made possible by the coinvestment of the WEST member community. This coinvestment enables access to a multifaceted tool that supports several WEST priorities, including facilitating and executing large parts of the biennial group collections analysis and supporting library collection management decisions with various reports and a robust on-demand collection comparison tool. Membership in WEST supports development and management of AGUA for significantly less than the cost of a bespoke collections analysis tool for each individual institution.

Every other year, when not conducting the biennial collections analysis, WEST undertakes systems development on AGUA, the WEST collection analysis and decision support tool. Planning for the seventh AGUA Development Phase began in the summer of 2024, with project ideas submitted by the WEST project team, the AGUA technical team, WEST Operations and Collections Council, and WEST membership. From the long list of possible enhancements, the OCC identified a short list of development priorities. The project team crafted a development plan based on OCC priorities and AGUA technical team capacity. The final development list and project plan were endorsed by the WEST Executive Committee in October 2024, and development work got underway in January.

This report provides an overview of the work completed in Phase 7, brief discussions of the benefits each project provides to members and the WEST program, and how the work fits into the larger contexts of the shared print and library communities.

WEST OCC

Kerri Goergen-Doll - Oregon State University
Wil Weston - San Diego State University
Cass Kvenild - University of Wyoming
Robin Sinn - Iowa State University
Leland Deeds - Washington University in St. Louis
Roger Smith - University of California, San Diego
Tim Kaneshiro - Getty Research Institute
Niamh Wallace - University of Arizona

AGUA team

Lucy Liu - Programmer Analyst
Joe Ferrie - Programmer Analyst
Dana Jemison - Metadata Analyst
Judy Dobry - Technical Team Manager
Alison Wohlers - WEST Program Manager
Nika Burns Teshin - WEST Collections Analyst

Overview of Phase 7 development activities

Planning for the seventh AGUA development phase (Phase 7) began in April 2024 and concluded in October 2024. Ideas for enhancement projects originated from a number of sources: the WEST project team, the AGUA technical team, the WEST Operations and Collections Council (OCC), and WEST members, who were invited to participate by submitting suggestions for enhancements that would help support local collection management workflows and decision making. Based on member and program priorities, the OCC identified a total of 17 projects to develop in Phase 7, the majority being small to medium backend enhancements to streamline functionality, increase the sustainability of AGUA, and enhance the data available through AGUA. However, the OCC also identified 3 high priority member-facing enhancements, one of which was a large-scale project aimed at enabling group filtering in AGUA, with the understanding that working on this enhancement might preclude the AGUA Development team from completing all 17 identified projects. The priority list of projects and development schedule were finalized in October 2024 by the WEST OCC, and development work began in January 2025.

In 2023, the WEST Executive Committee endorsed a recommendation to reduce AGUA systems development capacity for coding and building of new functionalities and tools by approximately half, targeting remaining capacity to support necessary AGUA systems changes and new, high-priority needs. This recommendation was affirmed by WEST members via a survey distributed to Directors and Primary Contacts in August 2023. This was the first development phase with this reduced systems development capacity, which meant there were fewer completed developments than in previous years. Although, the AGUA Technical Team still completed all of the enhancements deemed high priority by the WEST Operations and Collections Council, in addition to several key backend enhancements.

User-facing enhancements

Member reports

Over the past several years WEST has worked to add new data points to various collections analysis reports to enhance the group collections analysis and the on-the-fly reports available to members. In Phase 7, WEST completed two updates of this kind, one which primarily added new data points to the All Archived Titles Report, and another which improved the download experience for many of the reports available through AGUA.

All Archived Titles Report New Columns (Ref 4)
- **Summary of project**: This development added fields to the All Archived Titles Report and Archived Titles export from the On Demand Collection Comparison report that are useful for member and project team decision-making.
- **Outcomes**: Now in addition to the useful information already available in the All Archived Titles Report, members can also view, sort, and filter on validation information:
  - Validation action (583$a)
  - Validation level (583$i)
  - Completeness validation notes (583$z)
  - Condition validation notes (583$z)

  Furthermore, this development added the estimated volumes field, which was previously only a part of the unarchived output for the On Demand Comparison report, to both the archived output of the On Demand Comparison and Analytics reports, and the All Archived Titles Report.
- **Benefits**: This development makes it easier for members to determine which archived titles have been validated and at what level. Additionally, adding the estimated volume field to the All Archived Titles Report and the archived On Demand Comparison and Analytics Reports can help members understand the length of a given title or run, and help members prioritize titles for gap filling, estimate space savings for deselecting, and more.

- **WEST blog post:** https://cdlib.org/cdlinfo/2025/08/01/new-fields-available-in-the-west-all-archived-titles-report/

Update AGUA Report Output Format to Make Downloads More User Friendly (Ref 5)
- **Summary of project**: Previously the collection comparison reports, archiving commitments reports, and archiving exhibits (among others) would open a text file in a new tab when users clicked to download them. Users then had to left click and save these as .txt files on their computers before they could be used. This was an un-intuitive process for new users, and could lead to some formatting errors when users opened these files in Excel. So WEST sought to make it easier for users to download reports from AGUA.
- **Outcomes**: Since Excel is the most common way users interact with these files, the AGUA Development Team changed the way these files download, so that they will now automatically download as .xslx files.
- **Benefits**: Now, users can immediately open and begin working with these files straight from their downloads folder without having to convert them into another format. This streamlines the process of downloading and working with AGUA reports.
- **WEST blog post**: https://cdlib.org/cdlinfo/2025/03/14/download-agua-reports-more-easily/

Enabling Group Filtering in AGUA

The AGUA On Demand Comparison tool is highly unique in the shared print arena. Any WEST member may use it at any time to compare local journal collections against the WEST archives, journal collections archived for other shared print programs in the US and Canada, and WEST member unarchived data. To increase the value of this tool for members, the AGUA team undertook a large scale project to enable filtering based on predefined groups in the On Demand Comparison Report *and* add an entirely new On Demand Analytics Report which allows users to generate reports without using an input file.

Enable Group Scoping in AGUA (Ref 2)
- **Summary of project**: This development is an enhancement to enable libraries to compare against the holdings of pre-defined groups and generate reports without an input file.
- **Outcomes**: This development both adds a new "group" selection to the On Demand Comparison Report where members can select from the following pre-defined groups to compare their holdings to:
  - Colorado Alliance of Research Libraries
  - Consortium of Academic and Research Libraries in Illinois
  - Center for Research Libraries
  - California State University Libraries
  - Greater Western Library Alliance
  - HathiTrust
  - Orbis Cascade Alliance
  - Statewide California Electronic Library Consortium
  - UC Libraries
  - Western Regional Storage Trust

  This development also adds an entirely new On Demand Analytics Report tool which members can use to create reports without first providing an input file, scoped to one or more of the above groups.
- **Benefits**: This development enables members to leverage AGUA and the data in it for collections analysis and decision-making, while factoring in the multiple, often overlapping groups that our members are a part of. For example, members could use the new group scoping feature to generate a list of titles archived by WEST members who are also members of another group to prioritize for deselection knowing that they can make use of other existing resource sharing relationships when accessing these materials.

- **WEST blog post**: https://cdlib.org/cdlinfo/2025/08/01/new-group-selection-in-the-on-demand-comparison-report-and-new-on-demand-analytics-report/

Staff-facing and backend enhancements

Each development phase includes projects to support behind-the-scenes work in WEST. Phase 7 included 4 such projects, two of which focused on making the data management workflows more streamlined and sustainable, and the other two of which focused on enhancing the data which WEST provides to members through the various reports available in AGUA.

Data Management

Script for Archived Titles File Loading (Ref 1.c)
- **Summary of project**: The previous process for loading archived titles files was considerably time intensive for AGUA's metadata analyst. Additionally, there were potentially large time gaps between file loads, which could otherwise ideally be set up to run sequentially over night, or over the weekend.
- **Outcomes**: AGUA programmer analysts created a script which allows archived titles files to be loaded sequentially, with less intensive handling by the metadata analyst.
- **Benefits**: This development allows better utilization of the metadata analyst's time and streamlines workflows for archived title ingest.

CRL Validator Installation and Testing (Ref 1.g)
- **Summary of project**: Previously, AGUA was using an older version of the CRL Validator for analysis and reporting on unarchived and archived ingest files. The newer version contains more up to date functionality and options.
- **Outcomes**: The validation reports we produce using the CRL Validator have been updated to the latest version. Note that since this updated version of the new CRL Validation Report was implemented mid-disclosure ingest, some of the current reports may still be of the older version. All members who submit an unarchived title ingest file this fall will see the new version.
- **Benefits**: The new validation reports have the following updates:
  - Errors for a record have been consolidated into a single row. So, for example, if a record has both an OCLC number mismatch and an ISSN mismatch, you'll see both errors indicated in a single cell/row, rather than splitting these out into separate rows in multiple report sections.
  - Members will now see additional OCLC sourced information concerning the input record's OCLC match, which will will help in resolving errors locally.
  - Error categories, explanations, and other validation documentation will not have changed.

Data Enhancement

On Demand Comparison Report OCLC Number Matching Update (Ref 1.h)
- **Summary of project**: Previously, when contributors submitted data for an On Demand Comparison Report, the function worked by matching the input OCLC number with the latest copy/version of the primary number downloaded to our database. In order to accommodate any number within the primary plus merged (cross-reference) "set" of OCLC numbers, this process needed to be updated to allow for the input OCLC number to match with any one of the numbers within this set.
- **Outcomes**: The matching process for the On Demand Comparison Report was updated to allow for the input OCLC number to match with any of the numbers within the defined "sets". Where an input OCLC number is not available, or where the OCLC number doesn't match on anything, the ISSN in the input record is used for a secondary match point. No changes were made to this process.

- **Benefits**: AGUA now supports better matching when users submit data to the On Demand Comparison Report, through implementation of the change outlined above.

Location Values Field Length Increase (Ref 1.i)
- **Summary of project**: Where member submitted location labels are longer than 32 characters (current database field length), they are truncated, causing confusion for the contributor in outward facing AGUA functions where locations are displayed. In order to avoid confusion, this project was aimed at extending the current allowable length for location labels.
- **Outcomes**: The database field length for location has been updated to 128 characters from 32 characters, which should accommodate the longest string length.
- **Benefits**: Location labels are no longer truncated in AGUA when longer than 32 characters. In addition to eliminating confusion, this development also enables future planned development to hyphenate the location and sub-location codes in AGUA.

Looking forward to Phase 8

The goal for the eighth AGUA development phase will be to incorporate developments which will help to set AGUA on the path to further actualize the AGUA Vision Statement. The WEST OCC and EC are currently developing a roadmap for AGUA which will help define community expectations around the role AGUA serves for WEST and the shared print community at large, and how we get there. Planning for and execution of the eighth AGUA development phase will necessarily require new approaches to prioritization and decision making by the OCC and the WEST Project Team. As in previous development phases, WEST members will be invited to propose high-value projects that will create new efficiencies, surface new information, and create new modes of engagement with the program and participation in shared print activities.