An "Agreements Engine"
For the BTAA Shared Collection

A BIG Collection "Systems of Trust for the BTAA Shared Collection" White Paper

> *Agreements Engine: the system that contains an operable record of print items that individual libraries of the BTAA have designated for joint stewardship according to the policies and operational agreements for discovery, use, and persistence outlined in the Agreements for the Shared Collection.**
>
> *The registry, item matching, management rules, and indexing services that provide "network-level intelligence" about the print items designated for joint stewardship by every BTAA library*

December 2024

A product of the Agreements Engine (aka "SCSB Test") Pilot Team charged by the BIG Collection Steering Committee

"AGREEMENTS ENGINE" FOR THE BTAA SHARED COLLECTION
A BIG Collection "Systems of Trust for the BTAA Shared Collection" white paper

> *This white paper is one of several envisioned to describe the components of the distributed systems architecture for the Systems of Trust for Print Collections.*
>
> *For reference and context, see ["Building the Future of Collective Stewardship: The Big Ten Shared Collection White Paper"](#).*

The BIG Collection Paradigm Shift

The vision for the BIG Collection is to unite the separate collections of the Big Ten Academic Alliance Libraries into one collection, shared and fully networked. There are many consortia and institutions seeking to move toward shared print collections. This paper and its companions describe the approach of the BTAA Libraries and the choices we have made in our approach to this task.

BTAA is investigating middleware that can provide a central infrastructure service for the next phase of our Shared Print Collection building. The purpose of the Shared Print Collection is to provide a functional start toward the goal of uniting the collections of the Big Ten Academic Alliance into one collection, shared and fully networked. This aim represents a paradigm shift from the "current state" and the systems and infrastructure built out over the last several decades to support it. For the purposes of argument in this document, and to highlight the difference from the envisioned future state, the current state can be characterized as "distinct and independent "owners" which manage their collections based purely on local considerations and select a subset of their items to lend to others. To be clear, this paradigm shift does not represent a departure from the current state, but an organic evolution that builds on everything the libraries have accomplished together over sixty years of collaboration. It is only through the trust built over decades of trials, mutual learning, and shared successes that the BTAA libraries are prepared to take this next step into shared stewardship and access.

For the purposes of this paper, we will refer to the Shared Print Collection in distinction from the "total print holdings" of BTAA Libraries. The Shared Print Collection is one component of the BIG Collection which comprehends all content formats (digital and physical) as well as shared services, expertise, and infrastructures across all manner of domains. Note that for definitional purposes, we are not "creating" the BIG Collection - it already exists and has existed for 200 years.[1] We are simply entering into a phase of intentionally managing the entirety of holdings in

---

[1] The initial 2023 [high-level "total holdings" collection analysis](#) from the Gold Rush tool demonstrates the level of distinctiveness in the print collection that has been achieved through its first 200 years. Out of 22.7M unique print book titles, 62% are held by only one BTAA library, and only 0.1% are held by all fifteen. On the serials side, this pattern is even more pronounced: out of 1.3M unique print serials titles, 74% are held by only one library, while only 0.004% are held by all fifteen. This analysis is for the BTAA Libraries as of fall 2023 and does not include USC, UCLA, Oregon, and Washington, who will join BTAA in August 2024.

the Shared Collection as one collection, rather than individually managing distinct collections. We envision this as a process of designating items from their current *distinct* stewardship (local policies) to *shared* stewardship (jointly governed and jointly managed policies). The act of moving an item from distinct/local stewardship to shared stewardship will be referred to as "**designating** items for the shared collection." Designating items for the shared collection will remain at the discretion of each institution.

Scoping for the "Agreements Engine"

Once items are **designated as "shared"**, they become the joint responsibility of the entire consortium to steward together and will be governed by the agreements and practices that the consortium establishes for effective stewardship of the shared collection. In other words, a **"designated item"** becomes subject to the policy "Agreements" put in place by the BTAA Libraries as a whole (the consortium). The Agreements Engine is a registry for designated items that operates according to rules defined in the Agreements policies.

The shared collection is currently just over 500,000 print volumes; the vision is for this to grow to tens of millions. From this starting point, we do not know the pace at which the shared pool will grow and scale. However, we fully expect the process of converting from local to shared will take many years and incorporate new content as it is collected by members. From a systems or infrastructure services standpoint, the vision is that every volume committed will be immediately incorporated into shared stewardship and that a set of operational agreements governing committed shared items will be immediately applicable and operable for that item. This white paper will refer to the primary piece of middleware that we envision as the backbone for building the Shared Print Collection as the "Agreements Engine."

We are exploring whether and to what extent SCSB as built provides a starting point for our conceptual spec for the Agreements Engine to:
1. identify the gaps from current state to what we envision, and gain an assessment of resources (cost and time) that would be necessary to close those gaps;
2. Or, if SCSB could be used for our immediate needs while the features in our desired future state are built out;
3. Or, if a viable future could be that components of SCSB are split out as separate systems units (ReCAP is exploring the feasibility /desirability of this);
4. Or, whether repurposing SCSB is infeasible and what we are looking at would be a fresh build of a new system (with the as-built SCSB software either serving as a bridge or not as the case may be)

The purpose of the "agreements engine" is to make the shared collection with the agreements that govern stewardship and access for its items be **visible and machine readable** for BTAA members. The Agreements Engine provides a joint System of Record for agreements as well as a set of services that operate on top of the corpus of designated items (e.g. matching, management rules, and indexing services, which are described further in this paper).

"Shared Collection agreements" (SC agreements) can be defined as collective cross-institutional agreements to take actions to ensure that an item (or its equivalent) will continue to be available to the entirety of the Big Ten community for ongoing access. An **"SC agreement"** can be defined as having two parts: an **access component** and **persistence component:**

- **Access component:** consists of agreements to take action to ensure that an item (or its equivalent) is discoverable, accessible, and usable for the defined user community. Agreements supporting access could address:
    - access to metadata for discovery
    - fulfillment characteristics / levels of fulfillment,
    - "On demand" digitization or metadata enhancement (eg. requests by users)
    - Circulation status
- **Persistence component:** Agreements to address elements such as
    - Conditions of storage
    - closed stacks status
    - preservation actions
    - Digital surrogate creation / maintenance
    - metadata improvement
    - Repair, restoration, replacement
    - Avoiding withdrawal below of a minimum number of print copies

Example use story

The following use story is "imagined" rather than "actual" and is provided here simply for illustrative purposes. We anonymized this use case to "Library A" for the example, but it could be swapped out for any institution in the BTAA.

"Library A" of the Big Ten Academic Alliance would like to commit 100,000 books to the Shared Collection. Here's what happens:

- **LIBRARY A**
    - Through a local process, LIBRARY A compares the candidate books to the threshold characteristics of shared collection agreements defined through the "MOU for the Shared Collection" to ensure they are willing and able to meet both the access component and preservation component before committing the items to the Shared Collection
    - After local confirmation, Library A ensures that the appropriate metadata fields are complete to allow awareness of such components as location, condition of storage, circulation status, etc and registers those items in the Agreements Engine so that the agreements are
        - Visible to all other BTAA libraries
        - Subject to the operable automations that govern all items committed to the shared collection
- **AGREEMENTS ENGINE**
    - The agreements engine then takes over and:

- Matches the "designations" from LIBRARY A against the full database of registrations (e.g. against all commitments made by other BTAA libraries)
- Updates the matching count for titles and items committed to the shared collection
    - NOTE: there is no threshold or maximum for how many copies of a title can be registered.
- Maintains awareness of the full set of defined elements (condition of storage, digital surrogate, metadata quality, etc) for both the access and preservation components across all items of a given title
- Maintains awareness of on-demand requests from users for digitization or metadata enhancements
- provides alerts at various defined thresholds for both access and preservation actions defined under the agreed "persistence commitments" as well as user-initiated on demand requests. These alerts are passed on to the member library that stewards the item
- **LIBRARY A**
    - Receives alerts about access or preservation actions relating to its committed items, takes those actions, updates the relevant records accordingly, and submits the updates to the Commitments Engine
        - NOTE: "submitting updates" should be an automated process by which changes in records are already being submitted through "real-time" automated exports of changes in records

Conclusion

While this pilot investigated the technical infrastructure, its twin–the Shared Print Agreements and Procedures Working Group–explored the policy and procedure facet of building a jointly stewarded, distributed shared print collection. In Appendix B of its final report, the Agreements Working Group observed:

> Operationalizing BTAA shared print agreements and procedures will only be successful if member library staff have insight into the holdings at other BTAA institutions. At the simplest level, this will require a database that records the holdings data for materials that have been designated or committed by each institution for inclusion in the Shared Collection.

The report goes on to describe the "Functionality required by a Shared Print Agreements Engine" and provides an essential set of core functions that. The current white paper elaborates the outline of what the Agreements Engine is and what it needs to accomplish in order to provide a sound technical infrastructure for the next phase of building the distributed shared print collection. Along with the final report of the Agreements Engine Pilot and the final report of the Agreements and Procedures Working Group, it forms a brief guide for the group that takes up the next steps.