Sufia - Fedora - VIVO - Blacklight Interoperability (SufIVO /Sufitro)

The VIVO and Fedora communities have recognized their potential complementarity for some time, as is natural given the origins of both in the library community. These discussions have typically focused either on ways Fedora could complement a typical VIVO or in return on how VIVO could complement a typical Fedora repository.

- The Datastar project from 2008-2010 at Cornell illustrated the use case of Fedora complementing VIVO by providing a repository for storing datasets uploaded to Fedora through VIVO. A customized VIVO provided an editing and discovery interface for dataset metadata, while Fedora provided repository functions for the dataset itself, such as support for dataset download and versioning.
    - The Deep Carbon Observatory, led by Principal Investigator Peter Fox at RPI, uses VIVO as a data portal in concert with CKAN and Drupal, as described in Implementation of Open-World, Integrative, Transparent, Collaborative Science Data Platforms
- Institutional repositories such as UR Research at the University of Rochester sometimes feature pages for the researchers themselves, in addition to repository collections and resources, to support browsing by researcher and to provide an extra incentive for uploading content. While not powered by either Fedora or VIVO, UR Research illustrates another straightforward use case for a person-focused profile system complementing a repository.

With the rapid growth of The Hydra Project and the transition to Fedora 4 with its greater reliance on RDF there new opportunities for closer connections between VIVO, Hydra, and Fedora. The Sufia development community is presently adopting the Portland Common Data Model as a default model for organizing hierarchical collections and the associated works and files. This RDF model focuses on the structure of one or more digital collections but leaves open the treatment of additional metadata and/or relationships with entities not part of the collection(s) themselves, such as people, organizations, events, places, publishers, and subject headings. While Fedora 4's structure can accommodate arbitrary triples and anticipates coordination with a triplestore, the primarily focus and native application support focuses on on relationships among objects in the repository itself as opposed to arbitrary additional relationships to external entities. The gradual movement of collection metadata from string values to things (entities), and the creation of stable, linked-open-data URIs for entities of all types implies this broader need to create, manage, edit, and query open-ended RDF. This is where VIVO capabilities may best augment current Fedora and Hydra capabilities by providing not only a triplestore but the ability to create and edit the metadata conforming to any ontology, whether PCDM itself, the VIVO-ISF ontology for use cases involving researchers and publications, or ontologies supporting custom metadata, external entity linkages, or controlled vocabularies represented with stable URIs. Vitro by itself makes no assumption about the ontogy (ies) involved and in effect provides a blank slate that can be loaded with an ontology and support basic editing immediately and customized editing through configurable form and viewing templates; VIVO sits on top of Vitro and leverages the VIVO-ISF ontology with a number of pre-configured customized forms and views for publications, grants, and other complex entities.

The following figure illustrates the most likely points of connection that need to be explored using the low-level tools built into the Fedora, Hydra, and Vitro-VIVO development environments:
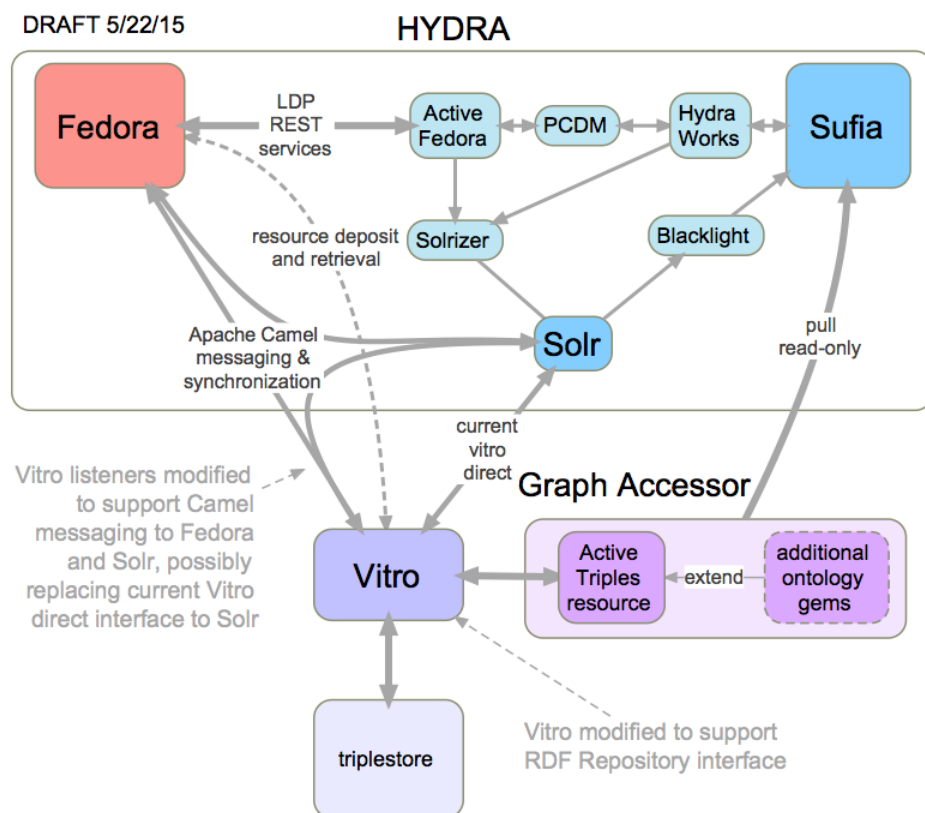
A number of platforms and communities could potentially be involved:

- Sufia (on GitHub) repository platform exemplified by Scholarsphere at Penn State

- - Try Sufia: A Fedora 4 Plus Hydra Combination
- Project Blacklight
- Apache Solr enterprise search platform, already used by both VIVO and Fedora
- Apache Camel routing and mediation platform used by Fedora 4 (see https://wiki.duraspace.org/x/mw8dB)
- W3C Linked Data Platform, also used by Fedora 4
- Linked Data for Libraries use cases
  - Use Case 1 (annotations to create virtual collections,tagging scholarly information resources to support reuse), both using Rails gems
    - ORE Ontology GEM implementation
    - Annotations as Linked Data with Fedora4 and Triannon
  - Use Case 2 (linking Cornell Library catalog with VIVO, URIs in library records), using VIVO
- ActiveTriples
- Hydra RDF Working Group

Use cases

- Agriknowledge use case -- add a flexible-metadata repository to a Blacklight app
- Climate Change Clearinghouse use case -- add a Sufia repository to a VIVO-Blacklight-Rails app
- Using Vitro and VIVO to extend the Fedora/based Sufia repository platform: the Sufitro/SufIVO use case