Lower the Bar to Contribution

- Recognition of many potential types of contribution – testing, sample data, documentation, training, ontology as well as code
- Clarification of the contribution process
- Simplification and modularity of code
  - Pushing configuration closer to the edge
  - Making visualization code more accessible to replacement and extension
- Improved documentation of code, process, workflow, best practices

What goes into a roadmap?

- Goals – where (all) are we going with VIVO?
- Opportunities – e.g., consortial search
- Strategy – how do we get there?
- Priorities – what do we do first?
- Players – who will be involved?
  - How can we leverage other efforts?
- Resources – how will it happen?

Roadmap process

- Writing user stories and estimating resource requirements
- Reviewing dependencies
- Defining appropriate chunks of work
- Voting
- Prioritizing and staging
- Committing funds and/or FTE

Likely implementation goals

- Streamlining each stage of starting a VIVO and ramping up to full scale production
  - E.g., reviving a virtual machine via Vagrant configuration and/or single-board computers
- Developing more robust ingest patterns
  - Making better use of common source data such as NIH RePORTER and PubMed
- Stronger liaisons with other open source communities to learn best practices and generate new ideas and approaches

Likely development goals

- Performance
- Modularity
- Configuration
- Search – SEO, relevance, local configuration
- Interoperability
  - Expanded terminology and authority services
  - Data flow from and to other tools – e.g., for visualization
- Responsive design & true mobile UI
- Provenance as in granular data review, error reporting, and logging/history

Architecture | Installation

- Modularity
  - Reorganize access to data models
  - Continued improvements to model-view-controller separation
- Distribution & installation
  - MySQL-free distribution (e.g., Jena TDB for 1.8)
  - Binary distribution to simplify deployment
  - Self-contained "instant" VIVO for pilot projects
  - Interactive configuration with embedded smoke tests

Performance

- Evaluate other triplestores using existing RDF API
  - Jena TDB, Virtuoso, AllegroGraph, others
- Redesign & reimplement simple reasoner to speed re-inferencing
- Restructure grouped property list
  - 60% of time required for profile page generation

Search within VIVO

- Unit tests with body of sample data to serve as verifiable baseline
- Support for additional facets
  - Contextual facets by type
- Configuration and tuning of relevance ranking
- Decoupling Solr to allow alternative search engine(s) like Elasticsearch or Funnelback (in 1.7)

Search engine optimization

- Dynamic sitemaps
- Better titles on pages
- Snippets with metatags
- Schema.org tags embedded in HTML to expose more of VIVO's internal semantic structure

Internationalization

- Further understanding our international users' needs
  - Occasional need, as for book titles in other languages
  - Interface and all content in a language other than English
  - Full bi-lingual or tri-lingual interface & content
- UI design to support multiple languages during content creation & editing
- Support for multiple languages in ontology editor and other admin functions
- Downloadable language bundles
  - Spanish in final preparation for use in Mexico, Costa Rica, Peru, and Spain
  - Active interest in Mandarin, German, Dutch, French

Integrity maintenance

- Better support for interactive and batch deletion
  - Removing all strictly dependent RDF
  - Leaving related authors, organizations, journals
- Site management data integrity tools
  - Scanning for orphaned data
  - Dead links
  - Being implemented as python/SPARQL tools at the University of Florida

Archival VIVO

- Maintaining historical record of former researchers
- Could provide authoritative information for external links
- How to separate out from current information in display, analysis, search
- Significant interface issues
- Should we run separate archival VIVO?

External URIs

- Extend VIVO support to new controlled vocabularies
  - E.g., Getty vocabularies being published as linked data
- ORCID integration through A&I grant
- Linking to people/organizations in another VIVO or via multi-site search index
- Linking to external organizational identifiers
- Coordination with library authority efforts, VIAF, & OCLC – in part through Linked Data for Libraries grant
- Modifications to better support persistence of VIVO identifiers and data

Ontology

- Tool to produce VIVO-ISF modules from broader ISF repository
  - Being implemented for eagle-i now
- Coordinated governance, evolution, and extensibility reaching well beyond the current VIVO community
- VIVO application ontology
  - Keeping ontology simpler while improving display & editing interfaces
- Ontology editor improvements
  - Unions and intersections for property domains and ranges

Ontology extensions

- Datasets and their relationships to publications, grants, projects, and contributors
  - Including information on downstream usage
- More detail on grants, contracts, & projects
- Impact, altmetrics, and usage
- "Knowledge Mobilization" & open government
- Facilities and equipment
- Library resources

Multi-site search

- Finish re-implementing linked data harvester from 2011 prototype
  - Harvesting in parallel, with interrupt/resume
  - Problem reports on harvested data
  - Mapping other ontologies to VIVO on harvest
- Refresh & extend front end of vivosearch.org
- Develop business model for hosting and participation
- Create and market disambiguation/resolution services using harvested data

Key Business Questions for Search

- Costs are reasonable but real, and have natural triggers like size & harvest frequency
- Operational challenges bringing new institutions and new platforms on board
- Versioning issues with ISF and changes going forward
- Other players in the market
- Great potential for building additional value
  - Lookup/disambiguation services on top of the search index
  - RDF mashups and distributed SPARQL queries

Alternative visions of VIVO

- A flagship product that needs to be easier to adopt, populate, and grow
- A vehicle for building networks of research data
- A reference implementation for the VIVO-ISF ontology as a standard for international data exchange
- A loose federation of many lightweight, creative apps
- A home and integration point for a virtual organization
- A tool to help universities and government agencies meet mandates for open government
- A discovery front end for repositories of documents & data
- These all exist and are not mutually exclusive!