|  |
| --- |
| VIVO: Enabling National Networking of Scientists |
| VIVO Functional Specifications |
| How VIVO should operate to the end user |

|  |
| --- |
| The Members of VIVO: Enabling National Networking of Scientists  5/19/2010 |

Table of Contents

[I. Background 3](#_Toc262643623)

[II. Summary 3](#_Toc262643624)

[III. Data Storage 3](#_Toc262643625)

[A. Scalability 3](#_Toc262643626)

[B. File management 3](#_Toc262643627)

[C. Security 3](#_Toc262643628)

[IV. Semantic Capabilities 3](#_Toc262643629)

[A. Searching 3](#_Toc262643630)

[B. Linking External Taxonomies 3](#_Toc262643631)

[C. Inter-ontology Mapping 3](#_Toc262643632)

[D. Ontology Editor 3](#_Toc262643633)

[V. National Network 3](#_Toc262643634)

[A. Aggregator Service 4](#_Toc262643635)

[B. National Network Search 4](#_Toc262643636)

[C. Linking External Individuals 4](#_Toc262643637)

[VI. User Interface 4](#_Toc262643638)

[A. Administrator 4](#_Toc262643639)

[i. Menu Management 4](#_Toc262643640)

[ii. Manual Data Ingest 4](#_Toc262643641)

[iii. Templating 4](#_Toc262643642)

[B. Individual User 4](#_Toc262643643)

[i. Data Review and Self Editing 4](#_Toc262643644)

[ii. Public/Private Data 4](#_Toc262643645)

[VII. Interface with External Systems 4](#_Toc262643646)

[A. Harvest Data 5](#_Toc262643647)

[B. Linked Data and Open SPARQL Endpoints 5](#_Toc262643648)

[C. Authentication 5](#_Toc262643649)

[VIII. Packaging 5](#_Toc262643650)

[IX. Risks 5](#_Toc262643651)

[X. FAQ’s 5](#_Toc262643652)

[XI. Glossary 5](#_Toc262643653)

# Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author(s)** | **Change Description** |
| 1 | 5/27/2009 | Christopher Barnes, Christopher Haines, Yang Li, Narayan Raum,  Dale Scheppler, Nicholas Skaggs, Stephen V. Williams | Document Created |

# I. Background

VIVO is a semantic web application built by Cornell in 2003.

# II. Summary

This functional specification will describe how VIVO will operate for the end user.

# III. Data Storage

## A. Scalability

## B. File management

## C. Security

# IV. Semantic Capabilities

## A. Searching

## B. Linking External Taxonomies

## C. Inter-ontology Mapping

## D. Ontology Editor

# V. National Network

## A. Aggregator Service

## B. National Network Search

## C. Linking External Individuals

# VI. User Interface

## A. Administrator

### **i. Menu Management**

### ii. Manual Data Ingest

### iii. Templating

## B. Individual User

### i. Data Review and Self Editing

### ii. Public/Private Data

# VII. Interface with External Systems

VIVO will interface with software and data from the institutional setting and the world. VIVO itself will function as a repository of ingested, generated, and hand entered data. This data will represent definitive information on the entities represented in VIVO. Harvesters will pull data automatically from the target sources and digest it for consumption by VIVO. VIVO itself will serve up RDF data via a web browser or RDF client. The data will be consumable in an easy and open way for users and developers alike. Linked Data and SPARQL end points will open the VIVO data store to the world wide web and allow VIVO data to be consumed by external sites and software.

## A. Harvest Data

The VIVO system will have the ability to ingest from authoritative sources, including human resources, grants, course databases, faculty reporting systems, personal citation management tools and web pages. Examples of data sources:

* PubMed
* NIH RePORTER
* Researcher.gov
* NSF Award Search
* CiteSeer
* PeopleSoft
* Sakai
* Blackboard
* Open Access Initiative Repositories
* Grantsfire Grants
* ISI
* Active Directory
* Google Scholar
* Scopus

Examples of data formats that the harvester will translate into the VIVO ontology’s schema:

* vCard
* hCard
* hResume
* hGrant
* EduPerson
* EduOrg
* Bibtex
* Medline
* OAI MetaData Harvester Standard
* NIH grant format
* NSF citation format
* NSF grant format

The VIVO system will ensure data quality remains high for all imported and entered data. All imported data will be standardized and checked for redundancy before import. The Harvester will be modular to allow system administrators to deploy the harvester in the best configuration for their institution.



## B. Linked Data and Open SPARQL Endpoints

The VIVO system will allow users to query for information from external systems. VIVO will provide two ways: native supporting linked data and instruction on installing plugins for a SPARQL end point. Examples of systems that will use linked data and/or SPARQL end points:

* Drupal
* Sakai



## C. Authentication

The VIVO system will support both native and external authentication processes. The external authentication process will support several known protocols:

* Shibboleth
* Kerberos
* Active Directory
* Internal Authentication

# VIII. Packaging

VIVO will include multiple methods for deployment including ANT builds and Virtual Machines.

# IX. Risks

# X. FAQ’s

# XI. Glossary

RDF

Linked Data

Shibboleth

Kerberos

Virtual Machines

Drupal

Sakai

Peoplesoft