

SOFTWARE METAPAPER

LIIS: A web-based system for culture collections and sample annotation

Matthew S. Forster¹, Lyn Paterson² and Robert J. Forster²

¹ University of Lethbridge, Lethbridge, AB, Canada, forster.matth@gmail.com

² Agriculture and Agri-Food Canada, Lethbridge, AB, Canada, lyn.paterson@agr.gc.ca, robert.forster@agr.gc.ca

The Lab Information Indexing System (LIIS) is a web-driven database application for laboratories looking to store their sample or culture metadata on a central server. The design was driven by a need to replace traditional paper storage with an easier to search format, and extend current spreadsheet storage methods. The system supports the import and export of CSV spreadsheets, and stores general metadata designed to complement the environmental packages provided by the Genomic Standards Consortium [1]. The goals of the LIIS are to simplify the storage and archival processes and to provide an easy to access library of laboratory annotations. The program will find utility in microbial ecology laboratories or any lab that needs to annotate samples/cultures.

Keywords: LIMS; database; CSV; archive; Sample Culture Annotation; metadata; metagenomics; Bootstrap; Codeigniter; PHP; HTML; MySQL; web-driven; MIMARKS

Funding Statement: Funding in support of this project was supplied by the Agriculture and Agri-Food Canada Genomics Research and Development Initiative (project # 2651).

(1) Overview

Introduction

A need exists for a simple tool to curate laboratory culture collections and handle ever increasing sample metadata annotations. As existing LIMS systems are complicated and cost prohibitive, the decision was made to create a lightweight, open source solution available to laboratories involved in culture collections or sample curation. The system would be a direct replacement of paper storage solutions and a parallel compliment to digital spreadsheets.

A design was proposed to index and organize laboratory data, help locate samples in freezers, and link any relevant information such as RNA/DNA isolations, trace and sequence information, culture pictures, taxonomy, animal and treatment information. Import and Export functions were added to further expand the utility of the program. Security is included within the system in the form of a password login and user permissions that can be applied to individual users.

The system is designed to have compliance to minimum information about a marker gene sequence (MIMARKS), a standard developed by the Genomic Standards Consortium (GSC) [1]. The sample context within the database stores all of the minimum information needed for submission to international sequence databases based on the GSC standards.

Started by a team at the Lethbridge College, the LIIS has gone through several development phases and has

been refined at the Lethbridge Research Center through a Co-op study program with the University of Lethbridge.

Implementation/architecture

Interface

A Web based interface hosted on a central server supports multiple users and geographic areas. This allows for labs to collaborate and access metadata from different locations, creating a central repository for projects with many participants. HTML views exclusively control web pages which allow for easy manipulation and extension of the application's output, with templates available to control aspects of every page.

Separate controllers for samples and cultures allows for conceptual and physical differences in the code, that grant future updates the opportunity to affect one context without changing the other.

User logins and permissions are included to increase the security and allow for limited access to the records. Permissions included can limit the user from adding, deleting or modifying records. This allows for a system owner to allow access to the records without worrying about unsolicited modification. A basic audit system is included within the tables which tracks the last modification of a record and who was responsible.

Focus on UI simplicity allows for easy use while saving time, which has been a goal of the project from the beginning. Simple design with singular purpose grants the application more meaningful use. Page design was based

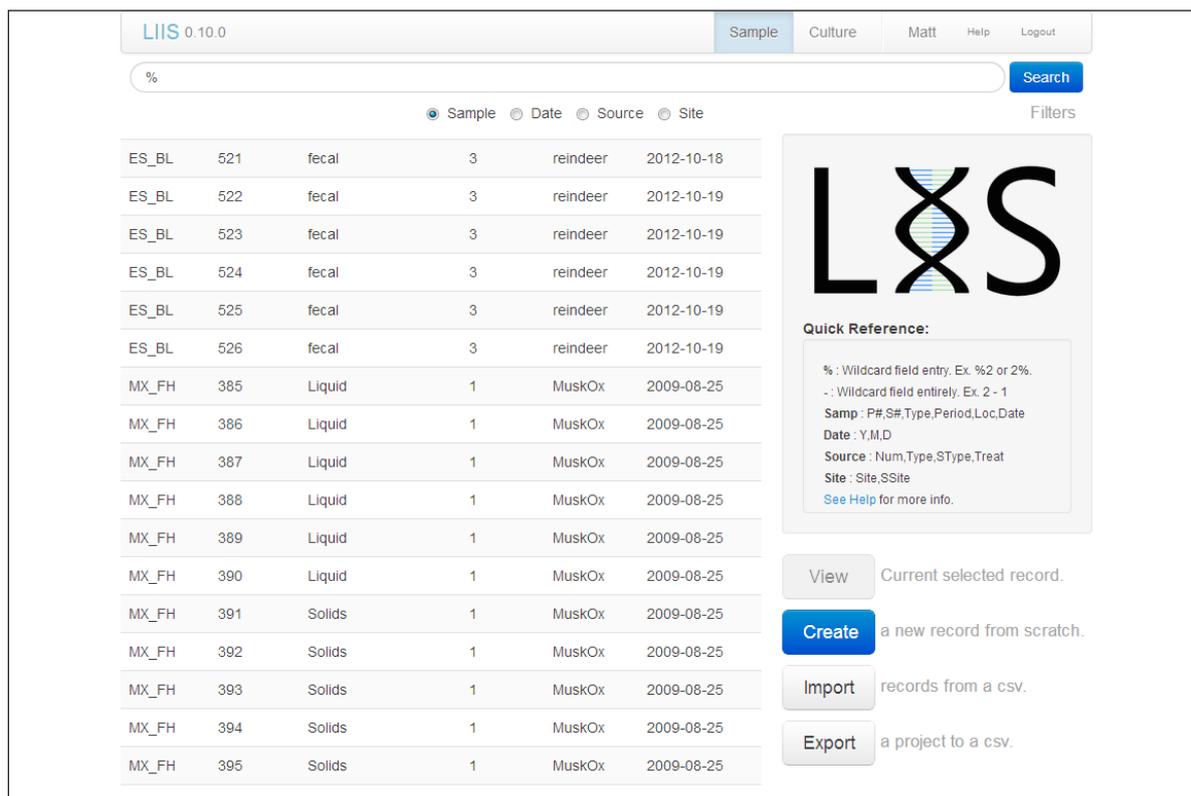


Fig. 1: LIIS Front page – search results.

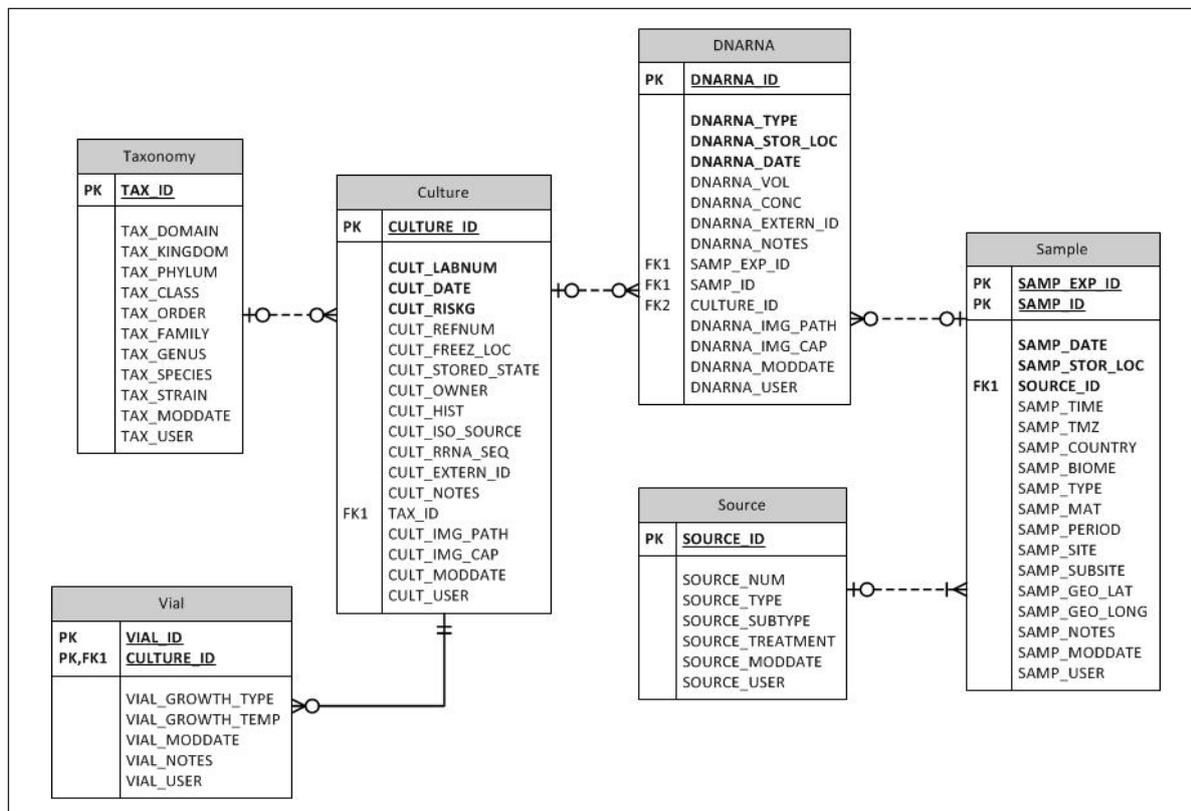


Fig. 2: LIIS Database Schema – Entity Relational Diagram’.

on Twitters' Bootstrap version 2.3.1 with jQuery 1.8.3 dictating the dynamic elements. Simple human-readable URLs and URIs allow for easy to understand manual navigation of the application.

An export function allows project samples to be exported in a CSV format using headers compliant to marker gene sequence (MIMARKS) standards [1]. This allows the easy selection of data for subsequent inputting into standard

data forms for the NCBI short read archive (<http://www.ncbi.nlm.nih.gov/sra>) or the European Bioinformatics Institute-European Nucleotide Archive (http://www.ebi.ac.uk/ena/about/sra_submissions).

Implementation

The Laboratory Information Indexing system was designed with the Code Igniter framework implementing the Model-View-Controller architecture. The abstraction of the database allows for easy extension/modification of pages as well as framework changes that do not affect the output.

The MySQL database (schema shown in **fig. 2**) has been de-normalized from a higher state to strike a balance between speed and disk space. There are two main tables with secondary lookup tables for frequently repeated or multiple case fields.

Distribution and installation

A Standard web server installation with a URL rewrite module available is sufficient to run the application, although the application was built using Apache and .htaccess files are provided within the application. Any web server capable of hosting Code Igniter can run the application. Basic knowledge about web server administration, and Code Igniter (and/or Model-View-Controller design) is needed for installation/modification.

A server hosting, or with access to, a MySQL service is required. The Code Igniter database configuration files (found in `application/config/database.php`) are required to be modified before the provided database installation script is run. MySQLi is used within the application, which requires MySQLi to be enabled within the systems PHP settings. Successful installations require medium-level MySQL and Apache/PHP administration knowledge, including modifying setting files (possibly both Apache and PHP), module activation and system privileges. A good reference guide for new webserver users can be found here: <http://ellislab.com/codeigniter/user-guide/installation/index.html>

- Ensure .htaccess is allowed in the target host directory (Apache AllowOverride option).
- Ensure the rewrite module is enabled in your web server preferences.
- Ensure MySQLi is enabled within PHP.
- Download and move the website files into the hosted directory.
- Special Write permissions are required for the resources/download and resources/upload directories, to ensure that the export and picture functions operate properly. Popular permissions for these folders are 775 for the appropriate Apache and PHP/FTP users.
- Open install.php and application/config/database.php within a text editor. Change the appropriate settings (MySQL username/password) to allow codeigniter to connect to MySQL. Make sure to change both the data and user database settings in database.php.
- In your browser, open (website_root)/install.php to install the database. If an error is shown, look over your

MySQL settings and the two files in step 5 to ensure connectivity.

- Log in to the user management with the user and password provided after successful database installation.

Proper permissions and related configuration is required for a working and secure installation.

Full instructions on installation can be found at: <http://www.github.com/forstermatth/LIIS>.

Information on Code Igniter can be found at: <http://codeigniter.com>

Quality control

LIIS has undergone functional testing in the form of Usability Testing, performed by the alpha development team during initial development of the interface. All database functions have been tested with up to 10 000 records for responsiveness and accuracy. All functions have been through basic unit testing during development.

Development versions are available through GitHub, and bug fixes are routinely addressed by the original developer, as well as being available to the community.

Testing was done on an Ubuntu Precise32 virtual machine with vagrant provisioning.

(2) Availability

Operating system

Any operating system that runs a web server efficiently.

Programming language

PHP5.3, JavaScript, HTML5, MySQL

Additional system requirements

None, other than the resources required to run the web server.

Dependencies

Not included: Apache 2.2+ (or equivalent), PHP 5.3+, MySQL 5.5.

Included: CodeIgniter 2.1.4, Parsedown 0.1.3, Twitter Bootstrap 2.3.2, jQuery 1.8.3.

List of contributors

- Matthew Forster – Lead Development, University of Lethbridge
- David Sinclair – Alpha Development, Lethbridge College
- Andrew Sanders – Alpha Development, Lethbridge College
- Graham Fluet – Alpha Development, Lethbridge College

Archive

Name

Figshare

Persistent identifier

<http://dx.doi.org/10.6084/m9.figshare.780836>

License

MIT (opensource.org/licenses/MIT)

Publisher

Matt Forster

Date published

23/08/2013

Code repository**Name**

GitHub

Identifier<https://github.com/forstermatth/LIIS>**License**MIT (opensource.org/licenses/MIT)**Publisher**

Matt Forster

Date published

23/08/2013

Language

English

(3) Reuse potential

The database fields have been designed for the most general cases, enabling use throughout many areas of study. The GSC standards [1] have been used where appropriate. The metadata stored within the LIIS complements the more specific environment packages curated by the GSC, and it is intended that they be used in tandem. These environment packages cover a wide range of applications, which increase the relevance of the initial metadata stored in the LIIS to a variety of scientific applications.

Modification of records is possible, where additional fields can be added / field names modified. Medium to advanced knowledge of PHP, HTML and MySQL is required to change the fields, and a high level of understanding of the code and which changes will affect what pages is required. Many of the base MySQL field names are built into the model code and should not be modified, but changing the name displayed by the HTML pages is easily done by modifying the HTML views. Adding fields requires modification of the MySQL tables, the input forms, the appropriate HTML view, and PHP controllers. Modification is only recommended to advanced users with a high level of knowledge about MySQL, PHP and HTML.

Both sample and culture contexts provide the opportunity for use in a variety of related and unrelated projects,

including laboratories involved in microbial ecology, culture collections, environmental microbiology, MetaHIT Consortium (Metagenomics of the Human Intestinal Tract consortium) projects (<http://www.metahit.eu/>), the NIH Human Microbiome Project (<http://commonfund.nih.gov/hmp/>) among others.

Examples of use

- Organizing and sorting culture and sample metadata in a time-insensitive format, providing opportunities for backup that would otherwise be unavailable to traditional paper or spreadsheet storage methods.
- A private, searchable metadata library accessible through as many clients as needed, with access to the library controlled through user logins.
- Quick access to specific storage locations and other information through a powerful field based search.
- The ability to annotate and review records from multiple computers, including adding Agilent Bioanalyzer traces, RNA/DNA quantification information and culture images.
- Convenient import and export of projects in CSV format. Exporting provides the data in a standards compliant format.
- Limiting access to create/update/delete with roles assigned to users.

Support

Support is continually provided for users by the active developer. General questions (installation etc.) can be directed to: Matt Forster (forster.matth@gmail.com, www.twitter.com/forstermatth).

Bug Reports and development additions/inquiries can be included through GitHub: [www.github.com/forstermatth/LIIS/issues](https://github.com/forstermatth/LIIS/issues)

Acknowledgements

- Stephen Graham - Supervising Instructor during Alpha Development, Lethbridge College
- Tim McAllister – Co-Op Sponsor, Agriculture and Agri-Food Canada, Lethbridge, AB

References

1. **Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, et al.** 2011 Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology*, 29: 415–420. DOI: <http://dx.doi.org/10.1038/nbt.1823>

How to cite this article: Forster, M S, Paterson, L and Forster, R J 2014 LIIS: A web-based system for culture collections and sample annotation. *Journal of Open Research Software*, 2(1): e2, DOI: <http://dx.doi.org/10.5334/jors.aj>

Published: 8 April 2014

Copyright: © 2014 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License (CC-BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/3.0/>.



The *Journal of Open Research Software* is a peer-reviewed open access journal published by Ubiquity Press

OPEN ACCESS The Open Access logo, which is a stylized 'a' inside a circle.