## 1. INTRODUCTION

## 1.1. Genomics and sequencing.

### 1.1.1. Sequencing data

### 1.2. Laboratory Information Management Systems (LIMS)

A LIMS is a software that allows the acquisition and management of the information generated in a laboratory. Until 1970’s management of laboratory samples and their associated analysis and reporting was a time-consuming manual processes so some organizations custom in-house solutions were developed by a few individual laboratories, while some enterprising entities at the same time sought to develop a more commercial reporting solution in the form of special instrument-based systems (1)⁠.

The first generation products appeared in the 80’s as the same time of the development of automatization of the laboratories. They were introduced in the form of a single centralized minicomputer, which offered laboratories the first opportunity to utilize automated reporting tools. A second generation of LIMS were were tapping into relational databases to expand LIMS into more application-specific territory. In the early 1990s, as personal computers became more powerful and prominent, a third generation of LIMS emerged. These new LIMS took advantage of the developing client/server architecture, allowing laboratories to implement better data processing and exchanges and processing of data anywhere on the network. Web-enabled LIMS were introduced the following years, enabling researchers to extend operations outside the confines of the laboratory. By the early 2010s, some LIMS had added additional characteristics that continued to shape how a LIMS was defined. Examples include the addition of clinical functionality and electronic laboratory notebook (ELN) functionality, as well a rise in the cloud-based software as a service (SaaS) distribution model (2)⁠.

Nowadays, their use has been expanded not only to sample and results management, but also to interface different systems and instruments. In addition, the importance of having a LIMS has increased due to quality regulations like ISO 17025 (3)⁠. Because of that, a certified quality laboratory must have a software to audit trail implementation, authentication protocols, configuration management, data backup, data integrity, electronic signature implementation, encryption, information privacy and network security (4)⁠. LIMS can contribute to it.

### 1.2.1. Existing LIMS.

We can find different LIMS vendors and they can be classified according to the way we can use them:

* Commercial brand LIMS: some specialized in automatization brands offer their own LIMS. That is the example of Illumina, which has developed specific LIMS for manage sequencing data.
* LIMS specialized software companies: there are several companies which develop LIMS and offer different services to adapt the system for laboratory necessities. A licence membership is mandatory to use this kind of software. We can find examples in different companies, such as LabWare, Thermo Fischer Scientific, Abbot Informatics (STARLIMS), etc.
* Open source: there are some LIMS open to all the community interested in implementing a software for managing laboratory data. This kind of LIMS are free and customizable, so is a low-cost alternative. Senaite/Bika LIMS, Free LIMS, Open LIMS are some of the examples.

### 1.2.2. Senaite-LIMS.

Senaite-LIMS is a web-based Open Source LIMS for enterprise environments, especially focused to behave with excellent performance and stability (5)⁠. It has launched on GitHub in October 2017 and deployed and maintained by two start-up companies, Naralabs (Barcelona, Spain) and Rinding Bytes (Nuremberg, Germany).

The software is a derivative work of another open source LIMS (Bika LIMS), built on top of Plone CMS with Python as its main programming language. The principle architectural changes with respect to its predecessor Bika LIMS is that they have developed Senaite as a system of independent add-ons. This, combined with quality code standards, makes the application much easier to maintain and to contribute to and allows it to stay level with emerging laboratory requirements. To help in achieving so, Senaite is developed under the paradigm of continuous integration (CI) and continuous delivery (CD).

## 1.3. Agile methodology.

## 2. METHODS

### 2.1. Project plan.

We followed the development of the system by agile methodology. A requirements matrix was made by a MoSCow chart. Futhermore, a Grant diagram was create to set project schedule, millestones and deliverables. We divided the project in three phases. In addition, a risks matrix was made to determine possible risks and their impact. All of these was transfer to JIRA software.

Finally, we created a GitHub repository to control versions of the objects and deliverables.

### 2.2. LIMS selection.

We selected Open Source Senaite-LIMS based on the next must-be criteria:

- Free and open source software.

- Customizable.

- Safety and back up.

- Trustworthy.

- Allow sample tracking.

- Identify who is using the application and what changes have done.

A search information about the different open source LIMS was done using the web page *LIMS Wiki*, which is a WIKI where we could find detailed information about laboratory systems.

### 2.2. **Installation**

First, we install an Ubuntu 20.04.1 64-bit operative system in Oracle Virtual Machine 6.1.14. Then, we installed Senaite 1.3.4 following the instructions of Senaite web page (6)⁠ and using virtual Python environment Miniconda.

As requirement of the software we use Plone 4.3.19, which is an open source and free content management system build on top of the Zope application server. Zope is a web application storage server written in Python language. The database management system used is ZODB (Zope Database) which is a database native to Python and makes it very fast for query or read operations because the language is not separate from it. In addition, ZODB provides a JSON-API for integrate Senaite and other systems. We use Python 2.7.18 and Zope 2.13.29.

### **2.3. LIMS configuration.**

After the installation we run the server instance and we configure the non-coding parts of the software. The configuration was made following the user manuals of GitHub Senaite repository and the information of Senaite web site.

The configuration was made to adapt the software to the necessities and workflow of genomic laboratory services. We use some real genomic laboratory services workflows and equipment of *Servicio de Genómica de la Universidad Complutense de Madrid* (7)⁠, *Laboratorio de Secuenciación del Instituto de Investigación Hospital Universitario La Paz* (8)⁠ and *Servicio de Secuenciación ADN, genómica y Proteómica de la Universidad de Salamanca* (9)⁠ and *Sanger Sequencing Service of Macrogen* (10)⁠.

- Laboratory departments, users and privileges and clients.

- Samples and analysis.

- Instruments and interfaces.

- Reports and notifications.

### **2.4. LIMS modifications.**

### **2.5. Functional tests and validation.**

In order to test the software workflows and configurations we made technical documentation to control changes and user manuals to understand the . All of these documents were copied and submitted on the GitHub repository.

## 3. RESULTS

### 3.1. **LIMS selection.**

Senaite LIMS was the chosen system, not only because satisfies the criteria detailed before, but also because there are extensive documentation and network support behind this brand. In addition, it has some additional characteristics that allows the developers and users to adapt the system to their necessities.

## **3.2. LIMS installation.**

We followed the steps detailed on Senaite web site to install LIMS in our operative system. In addition, we had to modify the rights of the created user (*senaite*). For this reason, we modified *sudoers.d* Ubuntufile, giving all rights to *senaite* user:

*#includedir /etc/sudoers.d*

*#Usuario actual*

*root ALL=(ALL:ALL) ALL*

*senaite ALL=(ALL:ALL) ALL*

## **3.2.1. Installation of the developed tool.**

## **3.3. LIMS configuration.**

#### **3.3.1. Users, roles and clients.**

To access into thesystem, three users have been configured (in addition to the admin role), all of them belong to the same laboratory: Sequencing Service.

- *Laboratory Manager*: this user can manage the LIMS from a functional perspective and have rights to modify analysis, instruments and also to verify and publish results and samples (user: labman / pass: labman).

- *Laboratory Clerk*: it can manage basic setup such as clients and it has rights to register samples and verify results (user: labclerk / pass: labclerk).

- *Laboratory Analyst*: it has no rights to manage, only to enter analysis results (user: labanalyst / pass: labanalyst).

Also, we created some clients in order to identify the samples:

*- North Hospital (NH).*

*- Public Health Institute (PHI).*

#### **3.3.2. Samples**

For samples registration we modified the fields of the “Request new analysis” leaving only the necessary ones.

The following Sample Types were also created:

- *Tissue*: for samples that requires previous DNA extraction and amplification (PCR).

- *DNA*: for samples that requires previous DNA amplification (PCR).

- *PCR Product*: samples that are ready for sequencing.

In addition, we modified the sample identifier (ID) to show the date of registration of the sample and the client.

#### **3.3.3. Analysis.**

We create the next *Analysis Categories:*

- *Pre-Sequencing*: analysis to do before sequencing process.

- *Standard Sequencing*: analysis belonging to the sequencing process itself.

The following *Analysis Services* were configured in to *Pre-Sequencing* category:

- *DNA extraction* (*nm ratio*).

- *PCR* (*pb*).

On the other hand, the next *Analysis Services* were configured into *Standard Sequencing* category:

*- Forward Sequence - GC content (%).*

*- Forward Sequence - Length (pb).*

*- Forward Sequence - Fasta File (-).*

*- Reverse Sequence - GC content (%).*

*- Reverse Sequence - Length (pb).*

*- Reverse Sequence - Fasta File (-).*

Furthermore, *Analysis Templates* were created for services to appear by default when samples are request:

*- Tissue – Bidirectional Sequencing:* to create samples with *DNA Extraction, PCR, Forward* and *Reverse Sequence* services and *Tissue* as *Sample Type.*

*- Tissue – Bidirectional Sequencing: t*o create samples with *DNA Extraction, PCR* and *Forward Sequence* services and *Tissue* as *Sample Type.*

*- DNA – Bidirectional Sequencing:* to create samples with *PCR, Forward and Reverse Sequence* services and *DNA* as *Sample Type.*

*- DNA – Unidirectional Sequencing:* to create samples with *PCR* and *Forward Sequence* services and *DNA* as *Sample Type*.

- *PCR Product – Bidirectional Sequencing*: to create samples with *Forward and Reverse Sequence* services and *PCR Product* as *Sample Type.*

*- PCR Product – Unidirectional Sequencing:* to create samples with *Forward Sequence* service and *PCR Product* as *Sample Type.*

**3.4. General configurations (*SetUp* menu)**

Senaite LIMS allows admin user some general configurations through *Setup* menu. For our case, the next modifications were done:

- Accounting -> *Include and display pricing information = False*

- Samples -> *Auto-receive samples = True*

- Analysis -> *Categorise analysis services = True*

*-> Default count of Sample to add = 1*

- Id Server -> We modify *Sample ID* field to show client, sample type and date information: {clientId}-{sampleType}-{yymmdd}

## BIBLIOGRAPHY

1. Gibbon GA. A brief history of LIMS. Lab Autom Inf Manag. 1996;32(1):1–5.

2. LIMSWiki. Laboratory information management system [Internet]. [cited 2020 Nov 2]. Available from: https://www.limswiki.org/index.php/Laboratory\_information\_management\_system

3. Standardization. IOBP (OBP). IO for. “ISO/IEC 17025:2017(en) - General requirements for the competence of testing and calibration laboratories”. 2017.

4. Wood S. Comprehensive Laboratory Informatics : A Multilayer Approach. 2007;(September).

5. Naralabs. SENAITE, Professional Open Source LIMS. The evolution of Bika LIMS. [Internet]. 2018 [cited 2020 Nov 11]. Available from: https://naralabs.com/en/blog/senaite-professional-open-source-lims-the-evolution-of-bika-lims

6. Senaite. Senaite LIMS installation [Internet]. [cited 2020 Oct 10]. Available from: https://www.senaite.com/docs/installation

7. Servicio de Genómica de la Universidad Complutense de Madrid. Genómica y Proteómica [Internet]. [cited 2020 Oct 5]. Available from: https://www.ucm.es/gyp/genomica-1

8. Laboratorio Secuenciación IdiPAZ. Guía práctica para el usuario de Secuenciación Sanger IdiPAZ [Internet]. [cited 2020 Feb 10]. Available from: https://www.idipaz.es/PaginaDinamica.aspx?IdPag=194&Lang=ES

9. Servicio de Secuenciación de la Universidad de Salamanca. Secuenciación de ADN [Internet]. [cited 2020 Oct 5]. Available from: https://nucleus.usal.es/es/secuenciacion/servicios

10. Macrogen. Standard Sequencing [Internet]. [cited 2020 Oct 5]. Available from: https://www.macrogen.com/en/business/ces\_service1.php