Baobab LIMS:

An open source biobank laboratory information management system for resource-limited settings

Hocine Bendou

Supervisor: Professor Alan Christoffels

South African National Bioinformatics Institute, Faculty of Natural Sciences

University of the Western Cape

This dissertation is submitted for the degree of

Doctor of Philosophy

SANBI March 2019

Declaration

I declare that ***Baobab LIMS: An open source biobank laboratory information management system for resource-limited settings*** is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Hocine Bendou

March 2019

Acknowledgements

I would like to express my appreciation to all those without whom this dissertation would not have been possible.

My deepest gratitude goes to my supervisor, Prof. Alan Christoffels for his valuable guidance, support, availability and consecrated time so that this study takes its merit.

I would like to extend my sincere appreciation to Dr. Dominique Anderson for the review of the content of the thesis, her advice and encouragements. I would also like to acknowledge the support of all SANBI members, the NSB Biobank personnel and the BIKA developers.

Also, my appreciations go to the National Research Foundation (NRF), the South African Medical Research Council (SAMRC) and the B3Africa initiative for the financial support of this project.

Finally, I would like to thank my family in France and Algeria, and my friends in South Africa, Belgium, France and Algeria for the support.

Abstract

A laboratory information management system (LIMS) is central to the informatics infrastructure that underlies biobanking activities. To date, a wide range of commercial and open source LIMS are available. The decision to opt for one LIMS over another is often influenced by the needs of the biobank clients and researchers, as well as available financial resources. However, to find a LIMS that incorporates all possible requirements of a biobank may often be a complicated endeavour. The need to implement biobank standard operation procedures as well as stimulate the use of standards for biobank data representation motivated the development of Baobab LIMS, an open source LIMS for Biobanking. Baobab LIMS comprises modules for biospecimen kit assembly, shipping of biospecimen kits, storage management, analysis requests, reporting, and invoicing. Baobab LIMS is based on the Plone web-content management framework, a server-client-based system, whereby the end user is able to access the system securely through the internet on a standard web browser, thereby eliminating the need for standalone installations on all machines.

The Baobab LIMS components were tested and evaluated in three human biobanks. The testing of the LIMS modules aided in the mapping of the biobanks requirements to the LIMS functionalities, and furthermore, it helped to reveal new user suggestions, such as the enhancement of the online documentation. The user suggestions are demonstrated to be important for both LIMS strengthen and biobank sustainability. Ultimately, the practi-

cal LIMS evaluations showed the ability of Boabab LIMS to be used in the management of human biobanks operations of relatively different biobanking workflows.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

`ACSR` AIDS Cancer Specimen Resource

`B3Africa` Bridging Biobanking and Biomedical Research across Europe and Africa

`BRIF` Bioresource Research Impact Factor

`GUI` Graphical User Interfaces

`H3Africa` Human Hereditary and Health in Africa

`ISBER` International Society for Biological Repositories

`IS` Information System

`IT` Information Technology

`LIMS` Laboratory Information Management System

`LMIC` Low and Middle Income Countries

`NIH` National Institute of Health

`NSB` National Health Laboratory Services Stellenbosch University Biobank

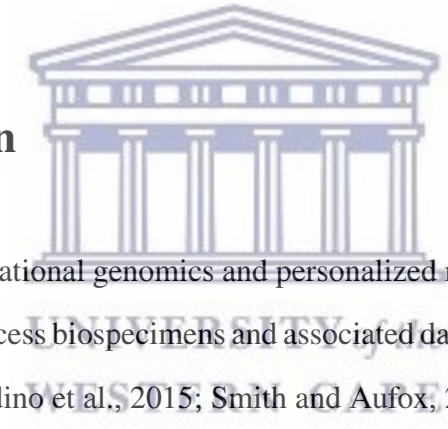
`QA` Quality Assurance

QC Quality Control

Chapter 1

Rationale

1.1 Introduction



Research projects in translational genomics and personalized medicine, irrespective of their size, essentially aim to process biospecimens and associated data in order to elucidate disease origin and etiology (Riondino et al., 2015; Smith and Aufox, 2013; Wilkinson et al., 2018). The outcomes and findings of combinatorial research studies, are then further explored by projects in other fields such as pharmacology and clinical trials, to develop and test therapeutics that may improve human healthcare (Matimba et al., 2008). Scientists and researchers, may for the purpose of their research projects, request biospecimens from biobanks, or collect these biospecimens themselves. Research groups which utilise the services of a biobank (which generally store high quality and regulatory body approved biospecimens), can considerably reduce research turnaround (time to result) which is generally impaired by difficulty and delays resulting from participant selection and recruitment, sample collection, and regulatory process approval (Sarojini et al., 2012). Biobanking operations are becoming more complex in terms of the continuously increasing quantity of the stored samples and

the addition of highly sophisticated biobank instruments with high throughput capabilities which have the potential to generate masses of non-human readable result files. Coupled to this, the regulations governing biobanks are subject to revisions and modifications, and best practices guidelines require continual refinement and adherence to maintain biospecimen chain-of-custody, and guarantee fit-for-purpose biospecimens (Bledsoe, 2017; Manders et al., 2018). As a consequence, manual methods can no longer handle the increased migration and variety of inter- and intra- laboratory information, and as such, automation of laboratory data management is becoming a necessity. Automated systems and instruments, enable biobanks to track and process a greater volume of biospecimens without need to hire and train additional staff. In addition, automation benefits to biobanks includes the elimination of errors, improvement of quality control, reduction in long-term costs and turnaround time, and ultimately, improvements in patient safety (DeLone and McLean, 1992; Berman, 2015). Automated systems include, but are not limited to, Laboratory Information Management Systems, Electronic Medical Records and Experimental Management Systems. Laboratory Information Management System (LIMS) applies computing technology for the integration of laboratory workflow, biospecimen lifecycle traceability, data collection from a variety of instruments and/or user input, notifications and reporting.

The Information Technology (IT) advancements to hardware performance, software development tools, and networking, all positively impact on LIMS development (McDowall, 1993). Modern systems tend to therefore demonstrate greater robustness, flexibility, and are easier to maintain. Many open source tools that provide valuable functionalities exist in public repositories and the integration of these functionalities allows for reduced time and cost associated with LIMS development. Open source tools are continuously maintained by developer communities, which contribute to the enhancement of the tools, while fixing bugs and attending to other issues within the source code. Wide Area Networking (WAN) technology has also influenced modern LIMS design, and newer information systems implement

a client-server architecture. Client-server based systems grant distant access to LIMS information, report and notification reception, and real-time interaction using modern Graphical User Interfaces (GUI) (Ulma and Schlabach, 2005). The implementation of LIMS can be a long and costly process (Rasmussen et al., 2007), as it requires (1) identification of all biobank operations and evaluation of automation ability, and whether their computerisation will improve biobank efficacy, (2) establishment of selection criteria and feasibility assessment of LIMS implementation, such as functionality lists, available budget, availability and skill level of laboratory IT personnel, and implementation deadlines, (3) evaluation of LIMS systems responding to the pre-established selection criteria. The evaluation verifies whether the system functionalities adhere to the biobank requirements, (4) allocation of physical resources for appropriate LIMS installation (Fig. 1.1). Moreover, workflows governing laboratory operations may be partially, or totally, discordant which makes it particularly challenging to find a "one-size-fits-all" LIMS that meet all biobank requirements (Rasmussen et al., 2007). Biobanks which are not able to make use of an existing LIMS may either; (1) adjust the selection criteria level, (2) request support and customisation from vendors, or (3) develop *de novo*, a tailored-made, in-house LIMS solution. The above-mentioned choices do present important drawbacks for the biobank's internal functioning. The first option may result in a reduction of automation levels which can reduce biobank efficacy, while the second and third options will require additional time, human and material resource investment, and may not fit with the selection criteria, timeline and budget (Kyobe et al., 2017). Biobanks in resource limited settings are generally small, less equipped, and may not have a skilled in-house IT team capable of developing and maintaining biobanking systems (Klingström et al., 2016). Sustainability in research biobanks, notably those operating in Low and Middle Income Countries (LMIC), is mostly based on short term funding from philanthropic organizations, external funders, governments, and other initiatives. Therefore, the feasibility of the aforementioned options may be further hindered (Kyobe et al., 2017).

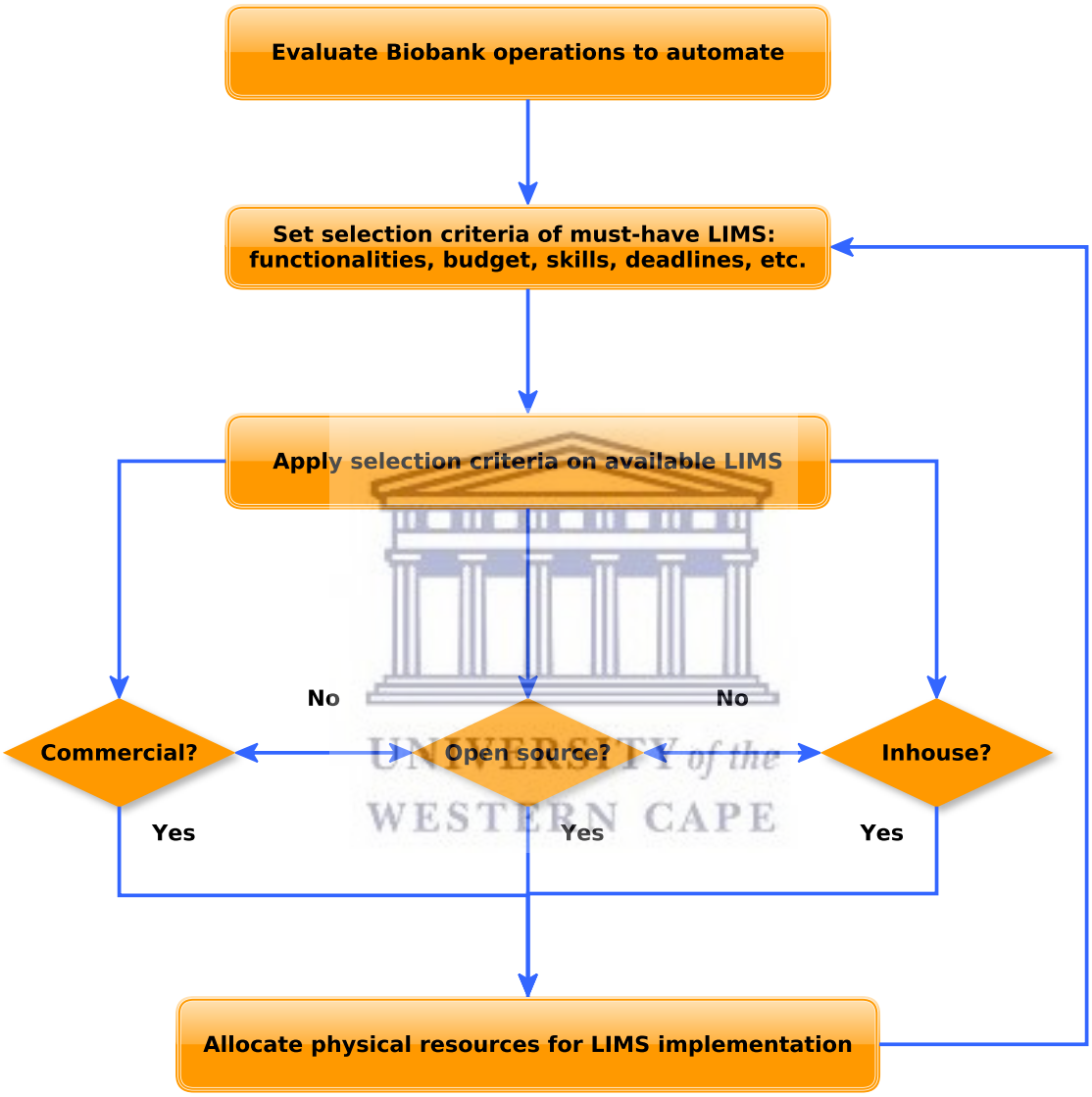


Fig. 1.1 LIMS selection plan. The selection criteria are inferred from existent laboratory resources and operations requiring automation, and are applied to the available LIMS products in the market; commercial, open source and in-house. LIMS systems responding to the criteria are tested and/or implemented.

There are a variety of commercial LIMS available on the market, however, they are extremely expensive. Licenses provided for commercial systems generally include a default set of functionalities or modules, but additional costs may accumulate for every additional customised requirement (Kyobe et al., 2017). Additionally, the headquarters of many commercial LIMS suppliers are based outside of LMIC zones, and as such, additional expenses may be incurred when the vendor specialists are required to travel to the biobank for physical implementation and/or face-to-face training (Kyobe et al., 2017). Unlike commercial LIMS, open source LIMS freely provide the totality of their source code in public repositories for use, sharing and customisation. Nonetheless, biobanks may demonstrate a reluctance in the implementation of an open source LIMS. This may be due to the opinion that open source LIMS are considered as less mature and present lack in functionalities, technical support and documentation, and standards and harmonisation (Piggee, 2008). Absence of technical support, standards and harmonisation appears to be of greatest concern to biobank stakeholders. Lack of support services can hamper biobank functioning if no other alternative solution exists. Furthermore, absence of standards and harmonisation in LIMS development makes data transfer difficult, via export/import functionalities, from one LIMS into another (Norlin et al., 2012).

1.2 Objectives

Biobanks which collect and store human biospecimens and related data, play an important role in the advancement of human healthcare through fostering research studies which aim to increase medical knowledge. Biobanks facilitate collaborations between different stakeholders from diverse backgrounds (donors, governments, scientific institutions, legal and ethical bodies, not-for-profit and for-profit organisations) to concentrate efforts in research while aiming to simultaneously strengthen biobank efficacy and governance (Yu, 2016).

Despite this, not all biobanks are benefiting from ongoing collaborative efforts and decision making. This may be due to the biobank's financial, geographical, and/ or political circumstances (Ciaburri et al., 2017).

LIMS implementation in a biobank can improve productivity, reduce costs, and aid in inter- and intra- biobank network formation (collaborations), and as such the LIMS can be an important factor in biobank sustainability. As stated previously, biobanks in resource-limited settings may not be in a financial position to employ a commercial LIMS, and available open source LIMS may not present as an attractive solution due to their possible inherent limitations. The study presented here demonstrates the development of Baobab LIMS; an open source LIMS developed for human population biobanks. Baobab LIMS was designed to deliver most of the pivotal functionalities required by state-of-the-art human biobanks, while remaining secure, scalable and easy to implement. Furthermore, we advocate its use in the LMIC setting, to overcome the limitations associated with commercial LIMS, and current open source LIMS.

1.3 Aims

The aims of this research project is to:

- Design and develop an open source LIMS for human biobanks. In this process, different technology solutions of open source LIMS were explored. The question of whether to create a *de novo* LIMS, or customise an existing open source LIMS was addressed
- Evaluate the LIMS using a real test case. The NHLS - NSB Biobank implemented the LIMS and was used as a case study for testing all of the integrated LIMS functionalities

- Enhance the LIMS by taking into consideration case study feedback and suggestions. Important motivations for enhancements were also obtained through implementing and interfacing of the LIMS within other biobank systems

1.4 Scope

The thesis is organised into five chapters. The rationale provides a general overview of the project and its importance in biobanking practices. The literature review (chapter 2) details the concepts, roles, and importance of biobanks, particularly in Africa. Moreover, a comparison between commercial and open source LIMS is used to motivate that an open source LIMS may be more suitable for biobanks in resource limited settings. Furthermore, some features of available open source LIMS are discussed in order to evaluate their usability in this research project. Chapter 3, outlines the computer science techniques employed in the design and implementation of the LIMS. The discussion and evaluation is reviewed in chapter 4, and the additional benefits of the LIMS is demonstrated in a use case. The project conclusions and further considerations are discussed in chapter 5.

Chapter 2

Background

2.1 Introduction

The explosive technological advances of the 21st century, particularly with respect to healthcare, has improved human quality of life and, in so doing, has increased life expectancy (Iskrov et al., 2017; Monteiro et al., 2018). The development of Computer, Robotic and Network technologies has led to the creation of services in the economic, financial and healthcare sectors, which are used in daily life and may be considered essential to human welfare (Joseph and Stone, 2003). Computational systems are present wherever management of information is needed, such as banks, markets, universities, hospitals, and laboratories, to cite a few. Services which were considered to be difficult and time consuming by both service providers and consumers alike, are today offered with less effort and in a shorter time. Instances which demonstrate the benefits of automation are numerous. For example, prior to automation, banks required the physical presence of the sender, followed by manual entry of the transaction information in a hard copy form. Manual checking of the information was a requirement to achieve a monetary transaction between two parties. These days, monetary transactions

can be done using online numerical services that are accessible to clients by using a computer linked to an internet network connection (Joseph and Stone, 2003). In terms of epidemiological studies in human healthcare, bioinformatic institutes within universities make use of the human reference genome, completed in 2003, to study the origins, causes and possible treatments for diseases which are a threat to human lives (Lander et al., 2001; International Human Genome Sequencing Consortium, 2004). This reference genome would not exist if a robotic sequencer connected to computer systems with immense hardware storage, was not developed (Lander et al., 2001). Prior to automation efforts, laboratory operations were tracked manually in logbooks, and experimental tests, results and reports were tediously, and often erroneously, written by hand (Prasad and Bodhe, 2012). Contemporary laboratories are now equipped with computer systems and Laboratory Information Management Systems (LIMS), that automate laboratory operations and generate legible and accurate reports using the push of a button (Prasad and Bodhe, 2012).

A recent innovation which has been listed among the top 10 ideas changing the world (Park, 2009; Kinkorová, 2016), is biobanking. This laboratory practice is playing an important role in the scientific community. Biobanks are considered to be a reservoir that supplies researchers with biospecimens (Riegman et al., 2008), the lack of which, can delay or even end valuable studies at the onset. Biospecimens which are collected, generally from preselected patients or populations, are intensively analysed in (amongst others) "-omics" and drug discovery studies to understand the etiology of the disease affecting the selected patients or populations (Smith and Aufox, 2013; Vora and Thacker, 2015). Understanding the intricate details of this revolutionary biobanking field, which is set to change biomedical research, allows for greater efficiency in biobanking practices by automation technologies, such as LIMS. In the sections which follow, the concept of a biobank will be further clarified.

2.2 Biobank Definition

The definition of a biobank has differed between organisations (Shaw et al., 2014), with the term "Biobank" being interchangeably used with that of a "Biorepository" (Marodin et al., 2012). Hewitt and Watson, (2013), in an effort to define a biobank, conducted a survey asking participants a series of general questions. Participants were individuals who were involved in the management of sample collections, and as such, would have existing knowledge so as to best define what a Biobank is. The general consensus in the survey showed that a Biobank is not only a sample collection system for human samples, but a system which could be employed for collection of a variety of different species and subspecies. The majority of respondents agreed that, to be called a Biobank, the management of sample collection must follow professional standards and that biospecimens must have associated data. Based on the combined findings, the authors proposed the following definition of the term Biobank (Hewitt and Watson, 2013);

> "A Biobank is a facility for the collection, preservation, storage and supply of biological samples and associated data, which follows standardized operating procedures and provides material for scientific and clinical use"

The definition above is, however, a general one, applicable to any kind of biobank and as such, does not further define the meaning of "biological materials and associated data". For example, with respect to biobank collections of human samples, the South African National Health Act defines a biological sample as; (Jordaan, 2016);

> "…material from a human being including DNA, RNA, blastomeres, polar bodies, cultured cells, embryos, gametes, progenitor stem cells, small tissue biopsies and growth factors from the same."

Associated data will include genetic, clinical, environmental and demographic information of participants and/or donors of biospecimens (Parodi, 2015; Harrell and Rothstein, 2016).

Therefore, for the purposes of this study, a human biobank is defined as;

> A facility for collection, preservation, storage, and supply of human biological materials including DNA, RNA, blastomeres, polar bodies, cultured cells, embryos, gametes, progenitor stem cells, small tissue biopsies and growth factors, and data derived therefrom comprising genetic, clinical, environmental and demographic information, which follows standardized operating procedures and provides material for scientific and clinical use.

2.3 Biobank Classification

In the last two decades, the complexity and utility of biobanks have expanded rapidly (Watson and Barnes, 2011; De Souza and Greenspan, 2013). Biobank capabilities have developed from small facilities intended to store residual biospecimens of single studies performed by single users, to large facilities comprised of several buildings (storage, laboratories, oversight bodies) which store and analyse a variety of biospecimen types. Modern biobanks follow strict standardized operations to preserve biospecimen and data integrity, and supervise access to data and biospecimens by ethical institutions to protect study participants (Watson and Barnes, 2011). The classification of biobanks assists stakeholders in determining adequate actions to be undertaken towards biobank growth and sustainability (Watson and Barnes, 2011; Hofman et al., 2013). For example, depending on the classification of a biobank, sponsors (funding bodies) may allocate varying levels of support (funds), researchers may or may not request data and biospecimens, and biobank managers, depending on the available resources, will be able to select the adequate system for laboratory information management and automation.

In literature, biobank classification was found to depend on (1) type of biospecimen they store e.g. tissue biobank (Campbell-Thompson et al., 2012), (2) specificity e.g. disease-based and population-based biobanks (Cervo et al., 2013), (3) scale i.e. small, medium and large biobanks (Roden et al., 2008) or (4) brand i.e. Commercial and Research biobanks (Anderlik, 2003). However, the classification criteria above, may not serve in the best interests of all biobanks and can hamper stakeholder activities related to the biobank. To this end, Watson and Barnes, (2011), proposed to categorize biobanks via a mono-, oligo- and poly-user system of classification, based rather on the number of users which the biobank is serving. Mono- and oligo-user biobanks are generally considered to be small scale and as such, tend to have less funding and resources when compared to poly-user biobanks (Watson and Barnes, 2011).

2.4 Biobank Importance

The main purpose for which biobanks have been established, is to empower scientific researchers with resources to use in their research endeavours (Doménech García and Cal Purriños, 2014). Biobanks provide (1) safe and secure storage facilities to maintain biospecimens for short, mid, and long term usage, (2) distant access via computer applications and/or catalogues linked to laboratory information management systems (to filter biospecimen and associated data content), (3) secure biospecimen supply and shipment services (Day and Stacey, 2008; Pang et al., 2017). In the post-genomic era, biobanks are oriented towards stratified medicine with the ultimate goal of precision medicine (Hewitt, 2011; Kinkorová, 2016; Merdad et al., 2017), that is, personalized diagnostics and tailored individualised therapeutics (Fig. 2.1). Despite the paramount number of publications demonstrating a reliance on these bio-resources, there is still a limited understanding of the importance and contribution of biobanks to medical research (Merdad

et al., 2017). To overcome this, initiatives which have been proposed, acknowledge biobank value by establishing a standard and clear citation of biobanks in published works which have utilised the resources of a biobank, called CoBRA: Citation of BioResources in journal Articles (Bravo et al., 2015). The Bioresource Research Impact Factor (BRIF) has subsequently been developed to quantify the impact and research outputs of biobanks as an alternative to the classification based on biobank user numbers, as stated above. This impact factor will be attributed to each biobank, in a manner similar to publication journals. As such, a biobank with a high impact factor may have a greater reliability and contribution in terms of biospecimens, when compared to one with lower impact factor (Bravo et al., 2015; Vora and Thacker, 2015).

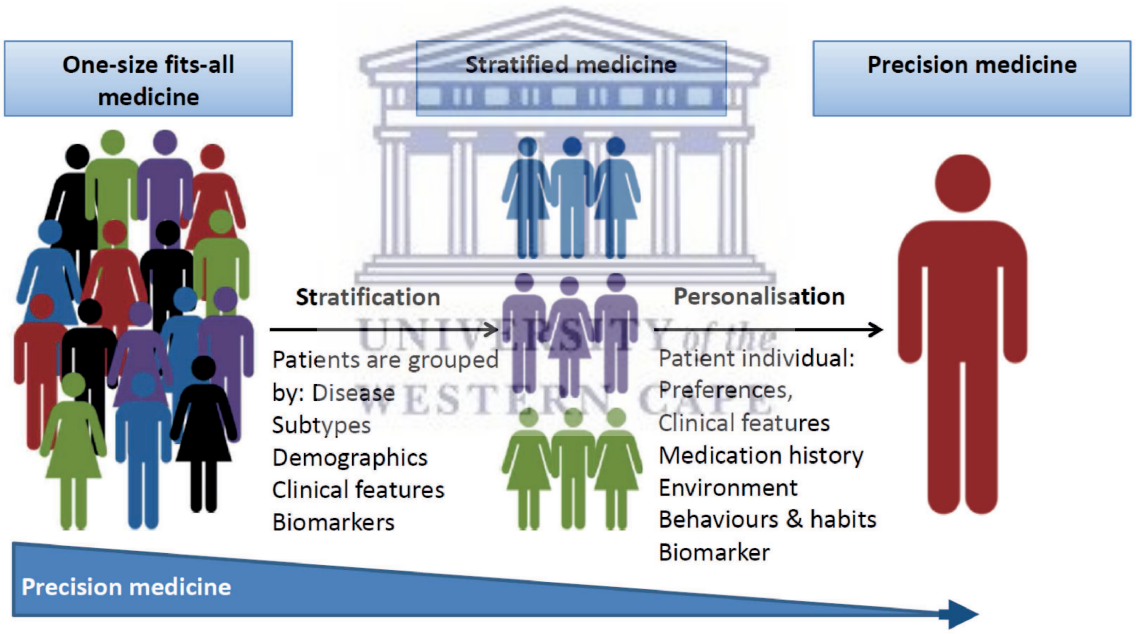


Fig. 2.1 Research evolution toward precision medicine. Precision medicine ensures offering of the right personalised medicinal care to the right patient at the right time (http://www.mpmi.manchester.ac.uk/aboutprecisionmedicine/).

2.5 Biobank Challenges

Biobanks are part of a complex network of diverse entities, each of whom have different visions, objectives and requirements (Fig. 2.2) (Kinkorová, 2016). The biobank entity is typically positioned at the bottom of the tree, where its internal functions are influenced by the fluctuating needs of the upper external structures. For example, a key requirement of a research unit is for samples of very high quality, while the primary objectives of a government and/or ethics body is the oversight and respect of a country's laws and regulations. External sponsors impact the usage and distribution of the biobank resources. Brief details of the important challenges which face biobanks is provided in the sections below.

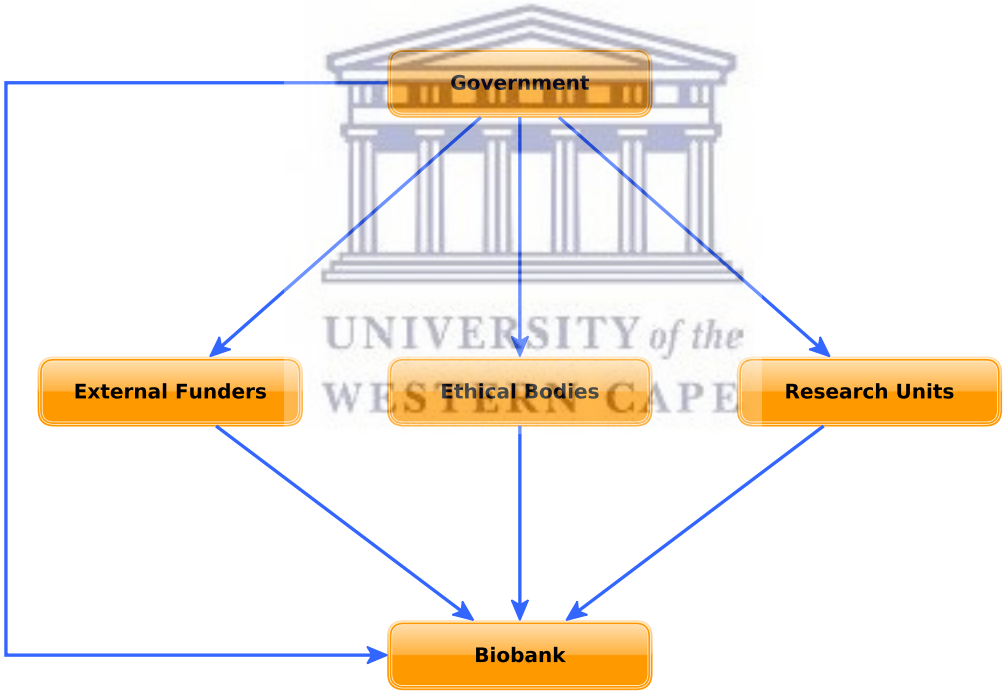


Fig. 2.2 Biobank governance from external bodies. Government and ethical bodies regulate biobanking operations, external funders set biobank objectives and research units require fit-for-purpose biospecimens.

2.5.1 Sample quality

The quality of biospecimens may impact positively, or negatively on the results of the translational phase of clinical and scientific research (Simeon-Dubach et al., 2012; De Cecco et al., 2009). It has been found that a substantial proportion of published research data is not reproducible (Moore et al., 2011). In USA alone more than 28 billion dollars are spent each year on pre-clinical research projects which are irreproducible (Freedman et al., 2015). In addition to the cost wasted on research and labour efforts, maintaining samples of inconsistent and/or low quality for undetermined periods before their identifications, presents an additional drain on biobank resources (Vaught and Lockhart, 2012). Differences in the pre-analytical operating procedures for handling samples within biobanks, has been implicated as the main reason for irreproducibility of research results (Ellervik and Vaught, 2015). In addition, combining result outputs obtained from samples of different origins may bias the final results in a collaborative research study (Ellervik and Vaught, 2015). International initiatives have been developed to standardize sample operating procedures, such as the International Society for Biological Repositories (ISBER) (Vaught and Lockhart, 2012; Vaught, 2016). The ISBER proposes guidelines and best practices related to standard operating procedures (SOPs) for collection, processing and archiving of biospecimens and associated data, and in this way, aims to promote harmonisation. Furthermore, ISBER accredits and certifies biobanks who comply with these guidelines and best practices (Betsou and Sobel, 2013). The ISBER guidelines are thoroughly reviewed and updated as new technologies and methods emerge, and current practices are revisited and consolidated with new SOPs to ensure "fit-for-purpose" biospecimens. Currently, ISBER best practices and guidelines is in its 4th edition and was released in the beginning of 2018 (Campbell et al., 2018). To further enhance development and adherence to regulatory requirements, societies/organizations in the biobanking sphere are being establish with dedicated and applicable standard guidelines for specific fields of research, such as the Canadian Tissue Repository Network Biobank

Certification and the College of American Pathologists Biorepository Accreditation Programs (Barnes et al., 2016; McCall et al., 2018).

2.5.2 Social, ethical and legal considerations

Research biobanks which utilise human subjects, or human biological materials and their derivatives, are subject to strict ethical regulations to ensure sensitive participant information are protected (Hansson, 2009). For example, the study of genetic information may lead to the elucidation of phenotypic abnormalities in participant groups, and could result in stereotyping or genetic exceptionalism (Green and Botkin, 2003; Vassy et al., 2015). As such, revealing genetic and health information publicly may harm participant integrity and due to the inheritance property of genetic information, harm can further extend to the participant's relatives and community, and result in stigmatisation (Green and Botkin, 2003). In addition, the repeated and suspicious use of participant health information will negatively influence the public trust in biobanks (Caulfield and Murdoch, 2017).

The protection of donors is the responsibility of ethics committees. The South African regulation Act No 61.2003, authorizes only health-related research studies approved by an ethics committee, to collect and analyse human samples. The ethics committee requires informed participant consent before approving a research proposal (Abayomi et al., 2013; Bolshete, 2015). The consent must (1) take into consideration and respect the participant's social and tradition beliefs, (2) clearly state the intention of the use of the samples with the exception of a broad consent in which samples can be reused in future-oriented research (Sheehan, 2011), (3) specify the possibility of return of research results and incidental findings to the consented participant (Appelbaum et al., 2014).

Furthermore, de-identification by anonymisation, or pseudonymisation, of the participant identity must occur. Anonymisation completely destroys the link between the individual and the sample information (Elger and Caplan, 2006). This destruction of the link between phenotypic and sample information, in certain studies, makes the interpretation of the results of a sample analysis difficult (Wallace, 2016). In contrast, pseudonymisation stores the participant, sample, and linked information in different storage databases (applications) with access granted only to authorized and trusted individuals (Strech et al., 2016).

2.5.3 sustainability

Sustainability is the capacity to "remain diverse and productive over time" (Watson et al., 2014). Biobanks mostly depend on external funding in order to sustain their operations, however, these funds may not be reliably guaranteed (Vaught et al., 2011; Henderson et al., 2015). Three pillars underpin sustainability; financial, operational, and social (Watson et al., 2014).

Financial dimension: encompasses three key areas; market strategy, customer focus, and brand recognition. From within these areas, the fundamental elements are the development and maintenance of a strategic business plan, and the cost recovery, generally by user fees.

Operational dimension: related to the (1) internal functioning of the biobank, that is, the operating practices followed to collect, process and maintain high quality biospecimens and associated data, (2) methods used to obtain clear and simplified consent forms, (3) mechanisms and systems in place to request data and biospecimens, and facilitate access to these.

Social dimension: encompasses two main aspects; community engagement, and accreditation and certification. The community engagement helps to educate, gain and retain public

trust, thereby enhancing donor enrolment. Accreditation and certification helps to gain trust of researchers, consumers and funders, thereby enhancing biobank utilisation.

The abovementioned dimensions cannot, however, be considered mutually exclusive of one another. In conjunction with general sustainability techniques, the efficient use of resources can reduce unnecessary expenditure and this cost reduction strategy is invaluable in biobank sustainability. For example, the number of biospecimens within a biobank accumulates over time and therefore decreases storage availability. Shea and co-workers, (2017), found that multiple freezers used in biorepositories are not optimally exploited, with 30% of their capacity unused. In this study, a storage assessment at the freezer and box levels helped to save 4055 boxes and 10.2 freezers in 23 collections (Shea et al., 2017). Similarly, adopting an open source technology for the automation of sample management information, may additionally economize the mostly excessive cost of laboratory information management system implementation (Kyobe et al., 2017).

2.6 Biobanking in Africa

Historically, Africa has been more greatly burdened by epidemic diseases, and the continent continues to register the highest prevalence of infectious diseases (Malaria, TB, HIV/AIDS) (Barr et al., 2016). More recently the Ebola virus caused thousands of deaths in Western Africa (Troncoso, 2015). Additionally, a noticeable increase in non-communicable disorders has been observed on the continent and it has been predicted that by 2030, 80% of cancer deaths will be in Low- and middle-income countries (LMIC), and Africa is expected to carry the major cancer burden (Mathers and Loncar, 2006; Ogunbiyi et al., 2016). In terms of infectious diseases, the expansion of the global transport network has the capacity to accelerate the spread of pathogens and their vectors and subsequently enlarges the infec-

tion boundaries within, and outside Africa (Tatem et al., 2006). Given this critical health situation, scientific communities have urged African national governments and international world organizations to take immediate action to prevent the intensification of health deterioration of African citizens. A review by Abayomi and co-workers, (2013), made note that the lack of biobank and bioinformatics infrastructures would negatively impact the ability for a rapid and efficient response to global outbreaks. The authors illustrated the urgent need to establish these two invaluable infrastructures locally in Africa, and close to the regions with high epidemic risks (Abayomi et al., 2013). In addition to considerations for the installation of physical infrastructure, the general lack of personnel skilled in biobanking and bioinformatics, must also be addressed, if the initiative is to be successful (Mulder et al., 2016).

In response, national governments and international organizations have noted warnings from the scientific community with regards to the dangers, and likely fatal consequences of global outbreaks. In the last decade, several biobanks were created in Africa through national and international collaboration initiatives such as the Human Hereditary and Health in Africa (H3Africa), the AIDS Cancer Specimen Resource (ACSR), and the Bridging Biobanking and Biomedical Research across Europe and Africa (B3Africa) projects (www.h3africa.org/consortium/projects, http://oham.cancer.gov/oham_research/programs/specimen_resource, www.b3africa.org). H3Africa for example, has funded the creation of three new biorepositories in South Africa, Nigeria, and Uganda. Furthermore, the H3ABioNet, a pan-African bioinformatics network for H3Africa, was established to support the development of bioinformatics capacity and promote scientific and biomedical research in Africa (www.h3abionet.org). The B3Africa project complements the work of H3Africa and aims to establish a harmonized ethical and legal framework for biospecimen and data sharing between the African and European continents. In addition, the B3Africa partnership is investing in the development of biobanking and bioinformatics open source systems and

tools, grouped in a ready to install virtual machine box (http://bibbox.bbmri-eric.eu). The work presented in this thesis, *Baobab LIMS*, is funded by the B3Africa project and was developed to manage biospecimens in biobanks operating in resource-limited countries (Bendou et al., 2017).

2.7 Laboratory Information Management System (LIMS)

An information system (IS) is defined as a set of components for collecting, storing, and processing data to provide information, knowledge, and digital products to support the management operations in an organization (Berisha-Shaqiri, 2014). Information systems are used by a variety of organizations (markets, laboratories) to perform and manage their operations, interact with their clients, and compete in the marketplace (Ranisavljević et al., 2012).

A Laboratory Information Management System (LIMS[1]) is an example of an IS mainly used in biobanks and laboratories (Olund et al., 2007). As biobank practices have evolved, so too have the requirement specifications for LIMS; from simple spreadsheets to highly sophisticated web-based client-server applications (Fig. 2.3). These adaptive information systems manage the information resulting from biobank operations and follow the whole lifecycle of a biospecimen, starting from the collection of the sample through to transport, processing, analysing, archiving and destruction of the sample (Fig. 2.4) (Moore et al., 2011; Redrup et al., 2016; Campbell et al., 2018).

[1] Also called BIMS for Biobank Information Management System

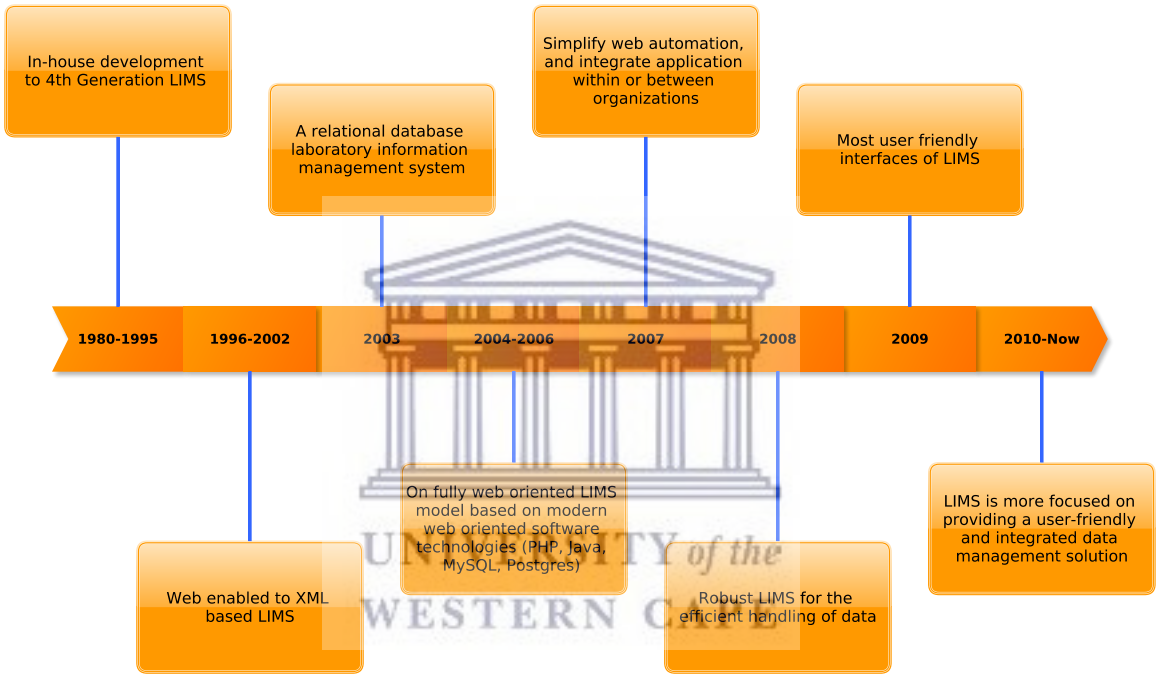


Fig. 2.3 Timeline outlining the evolution of LIMS. LIMS started as in-house development fulfilling requirements of a specific biobank, and evolved into sophisticated web-based application targeting wide range of biobanking workflows. Adapted from (Prasad and Bodhe, 2012).

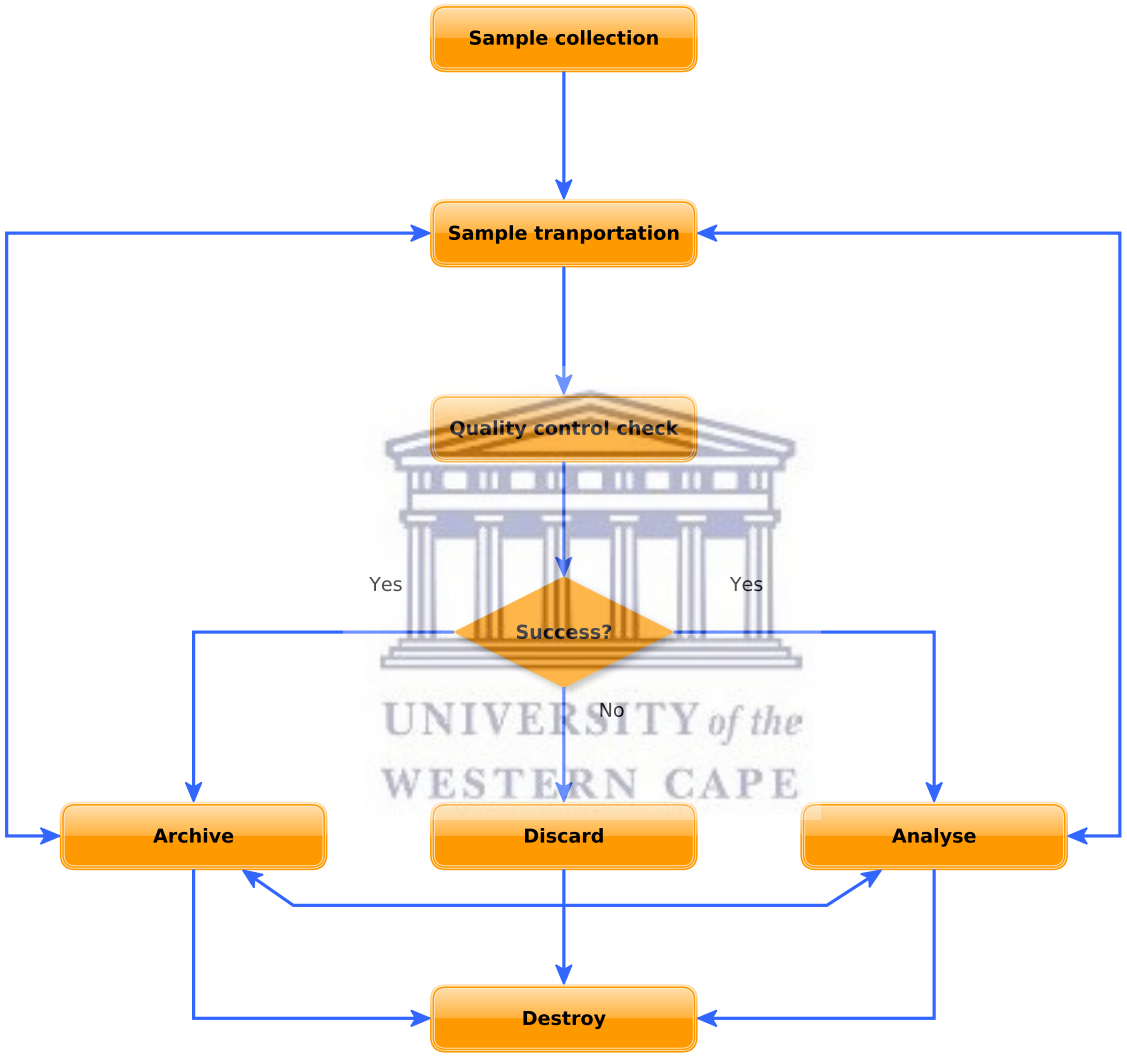


Fig. 2.4 Biospecimen life cycle workflow operations. Collected biospecimens undergo quality control analyses. Depending on quality obtained, biospecimens are either analysed, stored, or discarded and destroyed.

2.7.1 Sample collection

Generally, the collection of biospecimens is initiated within a research study targeting a specific disease, or a population behaviour (Capocasa et al., 2016). The collection of a sample may occur in the same vicinity hosting the biobank, for example in a hospital, or from a distant location, in the field, using ambulant laboratories. In some cases, a LIMS may foresee the type and the number of biospecimens expected to arrive at the biobank. In such cases, the biospecimens are created with unique barcodes, in the system prior to their arrival. Furthermore, certain biobanks may assemble kits containing empty barcoded tubes and ship them to the field where collection takes place. Upon receipt of biospecimens, information such as barcodes, date of collection and sample types are checked for consistency and logged into the LIMS.

2.7.2 Sample transport

Following collection, the biospecimens are transferred to the biobank via an internal, or external courier. Ideally, the LIMS should be able to track information about the date of shipment, date of arrival, and other pertinent information regarding the transportation of the sample (such as temperature, package condition, adherence to shipping instructions).

2.7.3 Sample processing

Once a biospecimen is received, the biobank may routinely perform a quality control check on the biospecimen. For example, if a whole blood sample is received and DNA is extracted, an analysis may be performed to check the concentration of the extracted DNA in order to verify that the biospecimen analyte has not degraded. Depending on the results of the QC

check, the biospecimen is either discarded, or processed further into aliquots and derivatives, and prepared for archiving. The result of the QC analyses, a log of the aliquots generated, and the biospecimen volumes should be tracked in the LIMS.

2.7.4 Sample archiving

The biospecimens which are vetted by the QC verification can be stored at an extreme low temperature in freezers and/or liquid nitrogen tanks (depending on the sample storage specifications). Exposing a sample to sudden freezing may undermine the structural integrity and stability of the sample components (cells, tissue) and as such, best practices for Cryopreservation and Biopreservation techniques are recommended (Baust et al., 2009). For example, there exists a cryopreservation method which allows direct transition of the biospecimen from liquid phase to glassy phase while avoiding the crystal formation phase that would result in cellular injury (Rall and Fahy, 1985). The exact storage location of the biospecimens and the temperature of the equipment in use for storage is critical information which must be tracked in the LIMS.

2.7.5 Sample analysis

The main objective of storing biospecimens is to use them in downstream analyses. The biobank may receive requests (generally from researchers) for biospecimens and/or their derivatives (aliquots). In response, the biobank curates the biospecimens in kits, in order to ship them to the defined destination. Large scale biobanks may have associated laboratories which are equipped with high throughput instruments, thereby enabling analyses to be performed within the biobank (Végvári et al., 2011). The LIMS should provide options to track the analyses which have been performed on a sample, the volume used, and the results

obtained. Furthermore, notifying clients by emailing them automatically generated reports which detail the analysis results, is a functionality that should ideally exist in the LIMS.

2.7.6 Sample destruction

Destruction of a specimen may be called for if one or more of the following circumstances occurs (Campbell et al., 2018);

- Low quality due to failure to adhere to pre-analytical procedures, equipment failure, or repeated freeze/thaw cycles
- Absence of associated data to recognize the identity of the specimen
- Error in consent, or participant withdrawal

The LIMS must therefore provide options to deactivate, or completely remove biospecimens and associated data from the system.

2.8 LIMS benefits

A LIMS may provide multiple benefits for all biobank stakeholders as it assists to;

Increase accuracy and speed of laboratory workflow process

Many biorepositories still rely on manual processes for recording laboratory information, despite the time constraints this imposes on the laboratory activities and the high possibility of inaccuracies due to human error (Magzumova, 2016). The drive behind laboratory

information management systems (LIMS) is to enhance biobank productivity and data quality (Dubey et al., 2012). As stated in ISBER best practices and guidelines, "the employment of a validated LIMS for automation of the biospecimen lifecycle is considered an obligation" (Campbell et al., 2012).

Manual entry of elaborate information, for example barcodes, increases the risk of transcription errors. With a LIMS in place, it is possible to rapidly scan and attribute a barcode to a biospecimen. Moreover, interfacing laboratory instruments with the LIMS allows for the automatic capture of high throughput data, and provides functions to verify the captured data, which therefore improves data quality assurance (McDowall, 1988).

Strengthen data integrity and safety

A LIMS provides functionalities to securely store data for long term periods, with the additional possibility to create replicates and back-ups of the entire data collection in different storage locations. Furthermore, details on changes in the captured data (insertions, updates and deletions) and the user responsible for the changes are automatically tracked in LIMS. This tracking of changes is particularly important in cases where data may have been compromised. In such cases, it must be possible to hold responsible person(s) to account and take corrective measures, such as a reversion to the original state.

Improve data accessibility and control

Previously, searching a sample in paper lab notebooks was similar to finding an English word in a dictionary. In a LIMS, a barcode will suffice to automatically fetch, almost instantly, all information about a biospecimen linked to the barcode provided. Additionally, it is possible

to access biobank information remotely, and securely, to further generate automatic statistical reports using the LIMS web capabilities.

2.9 Considerations for LIMS selection

An ideal LIMS will cover the majority of the biobank operations. However, to identify the ideal LIMS is both a complex and challenging process (Edward et al., 2008). Biobanks run different operational workflows and many of these workflows are not publicly available. As such there exists no standard LIMS that meets the requirements of all the existing workflows. Even if one application was created, the final LIMS product would be of incredible complexity which in its essence, would undermine the viability of the system. A premature and uninformed decision of which LIMS to use in the laboratory, may cause the automation objectives to fail and waste considerable time and budget investments (Mahaffey, 1990). The following considerations should be taken into account when choosing a LIMS;

- The system features must match the laboratory user requirements (Table 2.1)
- There should exist flexibility to accommodate evolving laboratory requirements. This may also have a positive impact on biobank sustainability
- Existence of demonstrable robustness of the system backend functionalities (database, backup system)
- The ease of use and availability of technical support, online documentation and training from the system vendor
- The budget allocated for the implementation of the system

Despite the existence of a system which meets the above considerations, laboratory personnel may exhibit a certain hesitance and/or resistance to change the laboratory operational protocols (Long, 1992). Additionally, employees may consider the automation process as a possible threat to career stability. Biobank managers should therefore make their personnel aware of the benefits of LIMS for the advancement of the laboratory, and involve them in the process of LIMS selection, especially considering that they may be more cognizant of the operations that require (or will benefit from) automation (Yu and Wilkerson, 2017). Laboratory personnel will be both positively and negatively influenced by the implementation of the LIMS. For example, while productivity may increase in the long term, adjustment and training may influence schedules and increase workload, in the short term (Avery et al., 2000)

Other specifications for LIMS selection have been outlined in Chapter 1 and additionally in the workflow illustrated in Fig. 1.1.

2.10 Commercial and open source LIMS comparison

One of the primary considerations to take into account following the decision to implement a LIMS, is whether to buy an off-the-shelf system, or develop a custom-built open source LIMS (Dermody et al., 2006). A comparative analysis of the pros and cons (as they relate to the operational dependencies of the biobank) will assist stakeholders in deciding which of the many available options to move forward with. There are differing opinions as to which version of a LIMS to implement and biobank managers may exhibit a preference for commercial LIMS over the open source technology (Avery et al., 2000), and vice versa (Bhandari and Snowdon, 2010). Table 2.2 highlights the differences and similarities between commercial and open source LIMS technologies.

Table 2.1 A list of typical LIMS features and requirements for the management of a human biobank.

Feature	Description
Sample tracking	Log information of operations performed on a sample and all states of the sample progression
Sample storage	Represent biobank storage as a hierarchy of sub storage levels (Freezer, shelf, box, positions) to precisely provide, and find a sample position in the storage
Inventory Management	Allow supply orders with a possibility to automatically adjust and track inventory storage quantities
Instrument integration	Interface with instrument to allow automatic upload of analysis results
Lab analyses	Log all information on analyses performed, samples used, analysts, and results
Reporting	Automatic client notification containing generated reports of analysis results
Invoicing	Automatic client billing and invoicing

Table 2.2 A feature comparison between open source and commercial LIMS products.

	Open source	Commercial
Initial cost	Free	Ranges in cost
Maturity	Lower quality and insufficient functionalities	Turnkey product with possible of lack in functionalities
Customisation	Source code available publicly for customisation	Customisation only possible by the vendor
Support & documentation	Generally, poor quality support and old or non-existing documentation	Excellent support service, with well written documentation, however, there may be an associated cost for this
Personnel	Community of developers with different levels of expertise and time dedicated for the project	Well structured organisation with full time skilled developers and managers

It is evident from the above comparison that commercial LIMS are more advanced than open source systems. Nevertheless, in view of the complexity and heterogeneity of biobank workflows, customisations are often necessary, regardless of the technology selected. Moreover, small customisations (that are in most cases associated to very high cost) are not always appealing to vendor services, and not all biobanks have the financial resources to entertain such customisation from the vendor. Resource limited biobanks cannot afford the expenses of a commercial LIMS and as such, open source LIMS technology is the best alternative solution (ARSLAN, 2014).

2.11 Open source LIMS examples

There are numerous open source LIMS available in the market, however, the majority of them were designed to manage very specific experimental data types, such as proteomics or Next Generation Sequencing (NGS) (Helsens et al., 2010; Scholtalbers et al., 2013). Therefore they cannot be used in biobanking automation. Below, a brief description of the two most well-known open source LIMS is provided;

2.11.1 OpenSpecimen

Previously known as caTissue, Openspecimen was initially developed under the caBIG program using U.S. National Cancer Institute funding (McIntosh et al., 2015). In 2011, the caBIG funding program ended and a community of developers from India took over the maintenance of the caTissue source code. In May 2014, caTissue was renamed to OpenSpecimen to underline the fact that caTissue is not only restricted to cancer diseases, and is an open source tool. From the official website (www.openspecimen.org), Openspecimen

is used in more than 65 laboratories worldwide. OpenSpecimen is a web-based application written in Java technology with a MySQL database management system used for the backend information storage. OpenSpecimen features and drawbacks are shown in Table 2.3.

Table 2.3 A list of implemented features and certain drawbacks of OpenSpecimen LIMS.

Features	Drawbacks
Sample tracking	Lack in functionalities – Laboratory instrument integration – Supplier orders and inventory storage – Kit assembly – Client analysis requests and services – Reporting of analysis results
Biospecimen storage	Source code complexity. Over 4 million source code lines
Biospecimen shipment	Expensive implementation and support services. From the official website, implementation is estimated at $45 000 US and support at $25 000 US per annum
APIs for integration with other third-party systems	
User authentication and authorisation	

2.11.2 Bika LIMS

Bika LIMS (www.bikalims.org) was developed in 2002 as a pilot, for management of laboratory information in the wine industry. The system has since seen a remarkable growth with major releases for water quality management and inter-laboratory proficiency testing. Bika LIMS is a web-based application written in Python, built with modern application server

and content management system employing the Zope and Plone frameworks. Bika uses ZODB, an object non-relational database, for information storage. Bika LIMS features and drawbacks are shown in Table 2.4.

Table 2.4 A list of implemented features and certain drawbacks of Bika LIMS.

Features	Drawbacks
Sample tracking	Lack in functionalities – Storage and freezer management – Supplier orders and inventory storage – Kit assembly – Creation of batch of biospecimens – Biospecimen shipment
Laboratory instrument integration	Has not been implemented in biobank management
Client analysis requests and services	Expensive implementation and support services
APIs for integration with other third-party systems	
User authentication and authorisation	

Bika LIMS incorporates a powerful laboratory analysis workflow (Fig. 2.5). Briefly, a client initiates a request for a sample analysis and a laboratory analyst performs the analysis on the sample and captures the results in the system. A lab manager must verify the results as a QC check and in the case where a result cannot be verified, it is rejected and tested again. At the end of the process an automatic report containing the analysis results is generated and sent to the client using an email notification system.

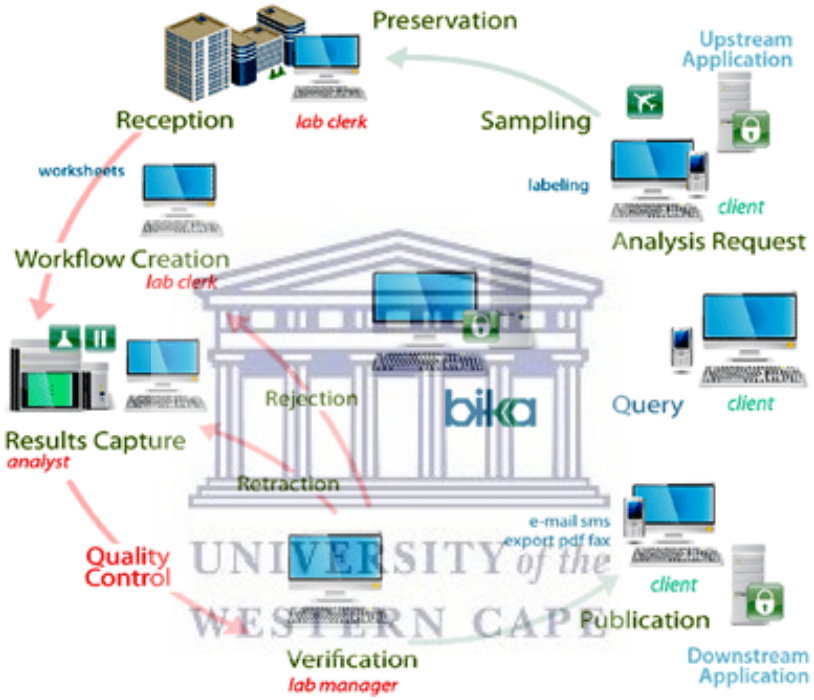


Fig. 2.5 An overview of the Bika LIMS analysis workflow. A client requests a biospecimen for analysis. The analyst performs the laboratory experiment and capture the results. The laboratory manager verifies the analysis results and forward them to the client in a report format (www.bikalims.org).

2.12 Summary

The resources and information kept within biobanks are necessary to enhance scientific research in a variety of fields. The management of these resources and the associated data is of vital importance to ensure consistency of the information and quality of the resources. There exist numerous laboratory information management systems (LIMS) capable of advancing laboratory productivity. Commercial LIMS are expensive to implement and to customise, whereas open source LIMS may lack functionalities, yet they are customisable due to the availability of their source code under licenses that provide rights for customisations. In the chapter which follows, the development of a LIMS with novel functionalities currently lacking in other existing open source biobanking solutions is discussed.

Chapter 3

Baobab LIMS: Development of an open source LIMS for human biobanking

Published paper

Hocine Bendou, Lunga Sizani, Tim Reid, Carmen Swanepoel, Toluwaleke Ademuyiwa, Roxana Merino-Martinez, Heimo Meuller, Akin Abayomi, Alan Christoffels

Biopreservation and Biobanking. Volume 15. Number 2, 2017. Mary Ann Liebert, Inc.
DOI: 10.1089/bio.2017.001

3.1 Abstract

A laboratory information management system (LIMS) is central to the informatics infrastructure that underlies biobanking activities. To date, a wide range of commercial and open source LIMSs are available and the decision to opt for one LIMS over another is often in-

fluenced by the needs of the biobank clients and researchers, as well as available financial resources. The Baobab LIMS was developed by customizing the Bika LIMS software (www.bikalims.org) to meet the requirements of biobanking best practices. The need to implement biobank standard operation procedures as well as stimulate the use of standards for biobank data representation motivated the implementation of Baobab LIMS, an open source LIMS for Biobanking. Baobab LIMS comprises modules for biospecimen kit assembly, shipping of biospecimen kits, storage management, analysis requests, reporting, and invoicing. The Baobab LIMS is based on the Plone web-content management framework. All the system requirements for Plone are applicable to Baobab LIMS, including the need for a server with at least 8 GB RAM and 40 GB hard disk space. Baobab LIMS is a server-client-based system, whereby the end user is able to access the system securely through the internet on a standard web browser, thereby eliminating the need for standalone installations on all machines.

3.2 Introduction

Human biobanking refers to the collection, processing, and storage of biospecimens and the collection of associated demographic and clinical data for future research use. The extensive collections of biospecimens throughout Africa collected for either specific research, population studies, or part of normal diagnostics workup were not necessarily collected for prospective use by researchers and practitioners. As a result, such collections might not necessarily have followed or adhered to evolving bioethical paradigms and international biobanking best practices (ISBER, 2012; Abayomi et al., 2013).

However, the establishment of the concept of centralized biobanks across Africa through initiatives such as H3Africa, the AIDS Cancer Specimen Resource (ACSR), and the B3Africa

(www.h3africa.org/consortium/projects, http://oham.cancer.gov/oham_research/programs/specimen_resource, www.b3africa.org) projects has highlighted the need for establishing and harmonizing national and regional biobank governance frameworks to address a relatively unregulated access to human and other ecological samples of academic interest in Africa. At the same time these governance frameworks fall in line with rapidly changing biobanking practices driven by modern technology (Dhai, 2013; de Vries et al., 2014) Similarly, a governance framework for IT infrastructure requirements that underlies a biobank does not exist yet.

According to the biological material tracking recommendations within the ISBER best practices, a computer-based inventory system should be in place to allow for the tracking and annotation of each incoming biospecimen into the biobank (ISBER, 2012). A laboratory information management system (LIMS) is thus central to the informatics infrastructure that underlies biobanking activities. To date, a wide range of commercial and open source LIMSs are available and the decision to opt for one LIMS over another is often influenced by the needs of the biobank clients and researchers, as well as available financial resources.

The National Health Laboratory Services (NHLS)—Stellenbosch University Biobank (NSB), a unit associated with the Division of Haematology at the Faculty of Medicine and Health Sciences, was established in 2012 initially through the ACSR project, and subsequently the NIH H3Africa funding initiative and required options for an LIMS implementation. The only option at the time was to consider a commercial LIMS because of time constraints to meet the growing need for biobanking services in South Africa. However, access to open source LIMS software allowed us to consider a longer term implementation that would align with our sustainability plans. Bika LIMS (www.bikalims.org) and CaTissue (now evolved and known as OpenSpecimen (www.openspecimen.org)) were identified as long-term options based on input from active software developer and user communities.

The Bika LIMS, although not specific for human biospecimens, is part of the BIKA software ecosystem that includes a BIKA Health for healthcare laboratories and Bika Interlab for interlaboratory proficiency testing. Customization of the Bika LIMS software would provide the benefit of inheriting a range of electronic health record functions that are central to establishing a core facility to support personalized medicine research. The recently funded European project, B3Africa, was established to strengthen IT infrastructure and ethical governance frameworks that would bridge biobanking and biomedical research across Europe and Africa. This funding provided the impetus to revisit the biobanking IT infrastructure at the NSB and to accelerate the development of Baobab LIMS, an open source LIMS for biobanking, as a strategy to provide a harmonized LIMS as an option for Africa.

A functional specification exercise (http://christoffels.sanbi.ac.za/index.php/projects/biobanking) in 2013 within the context of NSB biobanking requirements identified the following key modules as part of the extension to the existing Bika LIMS software, namely biospecimen kit assembly, shipping of biospecimen kits, storage management, analysis requests, reporting, and invoicing.

3.3 Implementation

Standard operating procedures (SOPs) associated with biological material inventory management were developed by the NSB team to inform LIMS workflow development. Other SOPs focusing on shipping, labeling, biospecimen procedures, and quality control were also developed in association with other H3Africa biobanks, and are publically available (http://h3africa.org/consortium/documents). The collection of SOPs underlies the NSB flowchart of biobank activities (Fig. 3.1) and subsequently the quality management system. Importantly, even with the use of electronic systems, it is important to keep hard copies

of all documentation detailing the biospecimen passage from reception throughout storage to dissemination as a QC check.

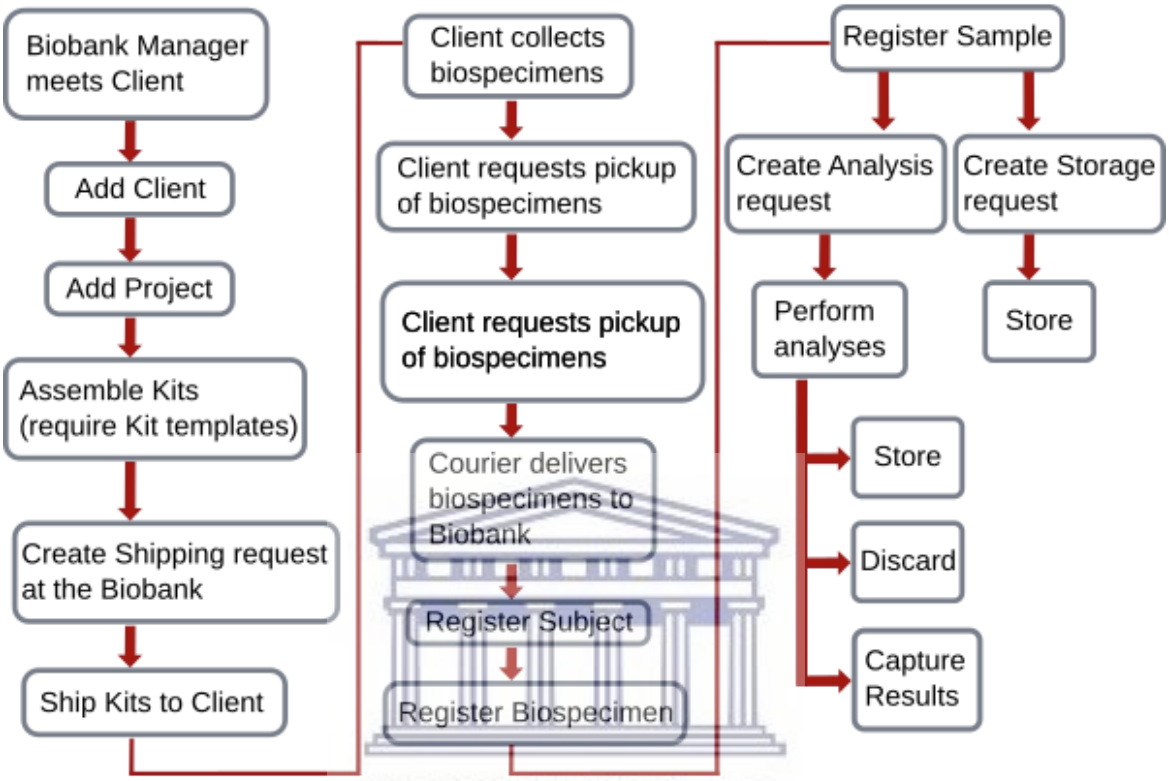


Fig. 3.1 Flowchart of the activities within the NSB biobank. The activities within the NSB Biobank starting with a face-to-face meeting with a client to the final storage of the biospecimen and/or the analyses associated with that biospecimen.

3.4 System Architecture

3.4.1 Hardware and software

Baobab LIMS is based on the Plone web-content management framework (https://plone.org). All the requirements for Plone are applicable to the Baobab LIMS. As such, the Baobab LIMS is a server–client-based system, whereby the end user is able to access the system

through the internet on a standard web browser, thereby eliminating disruptive workstation installations and inefficient software maintenance.

3.4.2 Workflow description

The following processes and tasks were carried out by the NSB and the client to ensure that a biobank project is successfully implemented (Fig. 3.1):

(a) The biobank laboratory manager meets with one or more client representatives to decide on the requirements of a new biobank project. A client does not refer to a physical person but rather to an institution such as a laboratory, a hospital, or even another biobank that can be represented by one or more representatives. In the Baobab LIMS system, they are called "Client Contacts." The project requires both parties to define the specimen types that will be collected and the biological analyses that are required at the biobank for all collected specimens. (b) The biobank prepares and assembles the kits that include equipment needed by the client to perform the sample collection, with one kit per participant. Mostly, a kit contains empty barcoded tubes and a form with a field for participant identifier, generated by the client. The collection forms are pre-barcoded with space for the client to use different barcodes for the tubes. In this way, the design of the forms minimizes the risk of assigning the incorrect tube to a participant. (c) The biobank ships the kits to the client by making use of external courier services. (d) The client receives the kits, collects the specimens in the associated tubes, and returns the kits to the biobank. (e) The biobank receives the kits from the client and registers the biospecimens in the system. (f) The biospecimens are then aliquoted for downstream analyses (the analyses defined in the project) or stored in freezers for later analysis.

3.4.3 Workflow implementation

The Boabab LIMS modules were designed to suit most laboratory specimen workflows, with minimal customization and postinstallation. In the NSB biobank, these modules were adapted to suit the specific needs as set out in the SOPs that were developed. These SOPs described in detail the life cycle of a specimen as it moves through the laboratory, and which steps require specific documentation, such as the time the samples arrived in the laboratory, the time at which sample accessioning took place, which user was performing the analysis, the method used in the analysis request (linked to its own SOP), and the time of completion of the task, to name a few. In this way, the LIMS was able to adapt to almost all of the manual documentation processes that occur in this specific workflow.

Client and analysis request modules were imported from Bika LIMS into Baobab LIMS and customized to accommodate human biospecimen laboratory services. The following modules were custom built to support the routine activities of NSB (Fig. 3.1).

3.4.4 Kit assembly

The biospecimen `Kit assembly` module provides the protocol needed to assemble kits that will be used to collect biospecimens in the field (Appendix B (Supplementary Data) and Supplementary Figs. B.1–B.3; Supplementary Data are available online at www.liebertpub.com/bio). This module ensures that clients can order the appropriate kits for their biospecimen collection. The project that is registered at NSB will define the appropriate collection kit(s). For example, if a project will collect samples for DNA and RNA extraction at the client's laboratory, then the appropriate collection tubes with an assigned function and label must be shipped to the client for immediate implementation in the field or for storage under appropriate conditions until it is needed. The NSB maintains an inventory of kit components

that can be put together and packaged for different sampling procedures and sample types depending on the client's need.

The kits include packaging material (styrofoam containers and corrugated shipping carton, absorbent material, and press lock biohazard plastic bag), gloves (optional), documentation (biospecimen submission form containing minimal information data, a shipping manifest (Supplementary Data and Supplementary Fig. B.4), shipping checklist, shipping query form, and a workflow on how to pack and ship the kit), and the specific biospecimen collection tubes that are already labeled. For example, a kit required for DNA and RNA sampling might contain a $6\mu L$ EDTA BD vacutainer or an 8.5 μL PaxGene Blood DNA tube and a 10 μL PaxGene Blood RNA tube or Tempus Blood RNA tubes depending on the downstream application. The kit assembly module provides the user the ability to select the appropriate kit template such as the templates for DNA or RNA sampling, which in turn is used to define the specific SOP required to assemble the material needed for the client's project. The kit template can be specific to one project or reused in different projects.

3.4.5 Shipping

The shipping module ensures that the correct instructions are given to send the appropriate biospecimen containers (as packaged in the kits) to the client and subsequent email notification to alert the client of incoming kits. Similarly, an e-mail notification based on the LIMS instructions will be sent to notify the NSB upon return of the kits from the client. The NSB, in consultation with the client, defines the appropriate containers to ship to the client and the shipping instructions to and from the client (Supplementary Data and Supplementary Figs. B.4–B.7). The kit, assembled in a size-appropriate box (Supplementary Data and Supplementary Fig. B.3), includes a manifest that describes each kit and the barcodes (Supplementary Data and Supplementary Fig. B.4).

3.4.6 Storage management

Biospecimen, kits, and stock items are handled within the inventory management module. The inventory management module describes the steps needed for the storage location so that products can be ordered and stocks updated accordingly (Supplementary Data and Supplementary Figs. B.8–B.14). Biospecimen storage covers a hierarchy of storage levels: storage unit (or room), freezers, shelves, and boxes. The biospecimens are stored in cryoboxes in various sizes depending on the size of the collection tube and cryotubes. For the other storage types (kits and stocks), there is no hierarchy to respect and the positions can be created at any level.

3.4.7 Freezer management

Different kinds of storage for biospecimens and aliquots exist inside the NSB depending on biospecimen type as well as the need for short- or long-term storage. The freezer management module describes the steps needed to define the structure that matches the physical storage in the NSB: rooms (within rooms), freezers, shelves, cryoboxes, and positions/locations (Supplementary Data and Supplementary Figs. B.15–B.17). A freezer contains shelves that contain cryoboxes. The last storage level at the NSB, namely cryoboxes, contains the positions reserved for biospecimens and aliquots. Similarly, the liquid nitrogen freezer and/or dewar contain, respectively, racks that contain canes. The box and/or cane can have multiple positions for biospecimen storage. Three classes (content types) were used to design the freezer management module (Supplementary Data and Supplementary Figs. B.15 and B.18), namely storage unit (room), storage level (freezer, shelf, and cryobox), and storage location (positions inside cryobox). This class inheritance was implemented using the object database ZODB (www.zodb.org/en/latest).

3.4.8 Security and administration management

Plone (the web-content management platform for Bika software) is based on the Zope framework. Zope provides built-in security functionality that allows us to define roles with permissions. A "permission" controls whether logged-in or anonymous users with a specific role can execute code and access contents. Each NSB staff member and client will have specific assigned roles and a level of security. Baobab LIMS also inherits Plone's secure version-controlled document management system, Plone workflow system and audit trails of back-end information changes made by the laboratory users.

Plone uses ZODB to store user data. ZODB, compared with SQL-based databases, is not vulnerable to injection as it uses binary format that cannot have user data inserted (www.plone.org). Plone authenticates users in its own database using a salted secure hash algorithm (`SSHA`) of their password. Using its modular authentication system, Plone can also authenticate users against common authentication systems such as `LDAP` as well as any other system for which a plugin is available (Gmail, OpenID, etc.). After authentication, Plone creates a session using an `SSHA` hash of a secret (token) stored on the server, the userid, and the current time. This is based on the Apache auth `tkt_cookie` format, but with a more secure hash function (www.plone.org).

3.4.9 Analysis request

A global list of available analyses is defined by the capabilities of the biobank laboratory. At the project level, analyses are defined for biospecimens based on the requirements of the project, such as DNA and/or RNA extraction applied to blood samples, and the resulting quality and purity results of the extracted DNA and/or RNA. Results of these analyses are registered and reported to the client (Supplementary Fig. B.20). The data are imported into

Baobab LIMS through an instrument interface such as the biodrop interface (Supplementary Data and Supplementary Fig. B.21).

3.5 Discussion

High-throughput genetic tools allow researchers to rapidly analyze thousands of biospecimens, as is common in consortia focused on population-based studies. In response, biobanks have to carefully consider the appropriate IT infrastructure that can meet the demands of large genetics studies. The Baobab LIMS was designed as an open source alternative for use in a resource-limited setting to meet the demands of increasing biospecimen collection.

During the development of Baobab LIMS, a new open source LIMS, Acquire, was published with a focus on pathology biospecimens (Dowst et al., 2015). This tool integrates the inventory management functionality of OpenSpecimen with modules specifically designed to meet requirements of researchers in a pathology laboratory. However, tools such as OpenSpecimen do not handle the specific biobank activities of NSB, which includes supporting prepackaging of biospecimen laboratory kits for clients (Fig. 3.1).

The establishment of centralized biobanks in Africa has drawn attention to the issues of interoperability between biobanks and harmonization of terminology used by each biobank. These concerns are not unique to African biobanks (ISBER, 2012). Data standards such as MIABIS (Norlin et al., 2012) provide an ideal platform to integrate data among biobanks. The advantage of using the same terminology across biobanks was demonstrated through the development of the virtual Breast Cancer Campaign Tissue Bank (Quinlan et al., 2015). Although a biobank catalogue is not currently available at NSB, we envisage a simplified application programming interface (API) that would allow users to access summary information using data stored in Baobab LIMS.

Quinlan and co-workers, (2015), suggested an increased role of ethical boards, governance, accreditation bodies, and funders to ensure that groups being authorized to collect samples have sufficient informatics capabilities to ensure the samples are used (Quinlan et al., 2015). This suggestion ensures that teams authorized to collect samples will also have the technical skills to ensure that the associated data are managed correctly. This concept relates to issues of biobank sustainability. Technical costs at a biobank go beyond supporting a LIMS and instead should incorporate a biobanking informatics management system that allows the biobank flexibility to meet the growing technical demands of making data available to a wider scientific community beyond biospecimen handling in a laboratory. In this context, and doubling the cost suggested by Dowst and co-workers, (2015), for maintaining their Acquire LIMS, the cost of employing a Linux administrator (20% full-time equivalent [FTE]), database administrator (10% FTE) and a computer programmer (25% FTE) in a South African setting would be 30,000 U.S. dollars (Dowst et al., 2015). The demands placed on IT staff require that biobanks have dedicated staff for these technical roles, thereby increasing personnel costs at least fivefold. These costs are a key consideration for centralizing and harmonizing biobanking in Africa. Unfortunately, local, regional, and international funders need to appreciate the importance of this critical component required for modern biobanking and the academic advantage of interoperability across evolving biobanks.

The adoption of Baobab LIMS by a wider user community will require more generic and configurable workflows. Nevertheless, Baobab LIMS has become a central component in the eB3Kit of the B3Africa project and is also included in the BBMRI-ERIC software catalogue, and in the open BIBBOX application store (Müller et al., 2017).

`Baobab LIMS website:` https://baobablims.org/

`Source code:` https://github.com/BaobabLims/baobab.lims.git

`Online documentation:` https://baobab-lims.readthedocs.io/en/latest

`Online demo:` http://baobab-demo1.sanbi.ac.za/. Use *admin* as username and password.

Chapter 4

Baobab LIMS. A practical evaluation exercise in a human biobank

4.1 Abstract

Laboratory Information Management Systems (LIMS) are valuable tools for automating and managing biobank operations. However, it is often a complicated endeavour to find a LIMS that incorporates and satisfies all possible workflow requirements that various biobanks may utilise. In this regards, software testing helps in the mapping of user requirements to the software functionalities, and furthermore, it may reveal new user enhancements and functionalities. The tracking of changes in the requirements early in the development stages helps to avoid development process delays and the exhausting of funds which, when they occur may subsequently result in project failure. In this study, we present a tangible practical test of the open source Baobab LIMS in a small scale biobank, namely NSB. The study demonstrates the importance of such an exercise for software enhancement and the influence of such enhancements on biobank sustainability.

4.2 Introduction

The Bridging Biobanking and Biomedical Research across Europe and Africa (www.b3africa.org) is an EU-funded program. The B3Africa initiative aims to facilitate scientific research collaborations between African and EU States, to provide an ethical and legal framework to enhance mutual exchange of biospecimens and associated data, and to release an informatics platform, the eB3Kit, containing a variety of biobanking and bioinformatics tools to support research activities (Slokenberga et al., 2017).

To meet the above objectives, 8 work packages (WP) were created (Fig. 4.1), each of which was assigned predefined project deliverables (Table 4.1).

Fig. 4.1 B3Africa work packages. Each of the 8 work packages was assigned predefined deliverables. WP3 was responsible for the implementation of an open source LIMS for the management of human biobanks (www.b3africa.org).

In addition to the mentioned work packages in Table 4.1, associate (beta) sites formed part of the development process of certain deliverables. The associate sites were established to assist in the design, conception and testing of concept aspects of certain deliverables work packages were responsible for. For example, for Baobab LIMS development,

Table 4.1 A brief description of the deliverables of each work package in the B3Africa project (www.b3africa.org).

Work package	Description
1	Provides a framework of legal and ethical regulations, which enable biobank biospecimen and data sharing
2	Creates the technical framework for integration of open source tools and hardware to be used in biobanking and bioinformatics research studies
3	Develops and provides a Laboratory Information Management System (LIMS) to be integrated into the platform (WP2) and used in the proof of concept (WP7). The Baobab LIMS is the final deliverable by the South African National Bioinformatics Institute (SANBI), responsible of the WP3
4	Provides the bioinformatics framework for the B3Africa project. It will provide guidelines for the implementation of bioinformatics analysis pipelines, use and reuse of experimental omics data, and sharing of datasets and results
5	Develops and implements a package for education and training on the use of the platform, best practices in biobanking, bioinformatics data analysis, and data sharing, observing ethics and regulations
6	Disseminates knowledge about the B3Africa project, the B3Africa informatics platform, and the possibilities it will provide for biobanks and research groups in Africa and Europe
7	Implements a platform as a proof of concept. Several partner institutions will take part as use cases to test the harmonization of ethics and regulations, biobanking best practices (LIMS in a Box), bioinformatics pipelines, and education and training in the B3Africa platform
8	Deals with the day-to-day project management issues (strategic, financial and legal, innovation, coordination of Work Packages, quality and risk management) by the Project Coordinator and the Project Manager

WP3 worked closely with the NSB situated within the National Health Laboratory Service (NHLS) and Stellenbosch University (SU) at Tygerberg Hospital. The NSB is an example of a growing biobank, storing human biospecimens, and initially established as part of the H3Africa project, funded by the US National Institute of Health (NIH) and the Wellcome Trust, to facilitate studies on Health, Disease and Pharmacogenomics of African Populations (https://www.sun.ac.za/english/faculty/healthsciences/haematological_pathology/biobank).

The NSB had initially not been using a LIMS to assist with the tracking of information and, as such, had evaluated OpenSpecimen and Bika LIMS. However, both systems were found to be unsuitable for their biobank operations. Baobab LIMS, an open source laboratory information management system, was therefore designed and developed by following the requirements and workflows of the NSB (see Chapter 3) (Bendou et al., 2017).

To this end, the workflow requirements of the NSB were translated into programming modules that constitute the core of the Baobab LIMS software, and the design of the modules was performed in conjunction with the biobanking team situated at NSB. Following development of each module, a series of tests and demonstrations were presented to the NSB team in order to evaluate whether the module functionalities met the specified requirements of the biobank. Each module was subsequently restructured according to the feedback obtained, retested and demonstrated to the NSB team. The cycle of feedback, software updates, and tests was repeated until no further recommendations were made, and the module functionalities met the biobank specifications. The development of the Baobab LIMS followed the waterfall methodology for the first version but due to the increased interest in the LIMS, the Baobab team has expanded to include a full time software developer and a project analyst, and the next version of the software will be developed using an Agile scrum methodology (Sliger, 2011). Sprints and milestones will be incorporated into the agile man-

agement system and will be invaluable in the development processes related to innovative requirements from biobanks.

System testing is described by the National Institute of Standards and Technology (NIST) as "the validation that the software meets its requirement. Validation of the complete system may involve many tests involving all system components. The software system tests exercise those system functions that invoke software to determine whether the software behaves as intended relative to complete system performance" (Klein, 2003). Furthermore, the Institute of Electrical and Electronic Engineers (IEEE) described acceptance testing as the operation that "checks the system behaviour against the customer's requirements (the 'contract'); the customers undertake (or specify) typical tasks to check their requirements" (Klein, 2003). The testing of any system is therefore vitally important to embark on, prior to the implementation of the system.

To investigate the ability of the developed Baobab LIMS to be used in automation and tracking of human biobank operations, the final product was implemented at the NSB biobank and the full functionalities were tested and evaluated in a practical exercise.

4.3 Methods

a. Test scenario

The organisational structure of the laboratory-based personnel working at NSB comprises of a lab manager, an IT system administrator, a lab clerk and two analysts. Throughout this exercise each user was assigned specific roles and tasks in the LIMS (Table 4.2). A laboratory in the Division of Human Genetics at the University of Cape Town is one of the NSB clients. The laboratory client accepted to participate in the LIMS evaluation, and a project

entitled "*Baobab test case*" was agreed between both the lab managers. The client, for the purpose of this defined project, is assumed responsible for the collection of whole blood specimen from two "mock" participants. It is important to note that the intention of this exercise was to evaluate the ability of Baobab LIMS to automate and track the information of the biobank and client operations. Therefore, to facilitate the progression of the exercise and avoid ethical and legal constraints regarding use of human specimens, the collection of whole blood biospecimens were replaced with other non-human substances. Consequently, there was no physical testing of the quality of the biospecimens. However, the information was entered into the LIMS in order to ensure that tracking of analysis would have occurred and been recorded appropriately, had the biospecimens been of human origin. The biobank assembled two kits, one kit per participant, each containing two barcoded 'BD EDTA' vacutainer tubes (Fig. 4.2(a)). The assembly process for each kit utilised products from the inventory storage. Both kits were packed according to IATA regulations and enclosed in a secure box, and subsequently shipped to the client using the DHL courier services (Fig. 4.2(b) and 4.2(c)). The client received the shipment. The standard workflow involves the shipment received by the client and subsequent collection of biospecimens from consented research participants. Thereafter, the biospecimens are dispatched back to the biobank using one of two possible methods; either by notifying the biobank to send a preferred courier to fetch the kits, or use the resources of the client for the dispatch. In the case of this exercise with non-human substances, the former option of dispatch back to the biobank was utilised following email notification sent from the client, to arrange for shipment collection. The NSB then received the kits and ascertained whether all vacutainers and corresponding documentation are in place. As a standard, for the assigned workflow, DNA extraction analysis followed by DNA purity and concentration assessment, as well as functional PCR would be performed to determine the quality of the biospecimens. In the case of this exercise, dummy

data were used as results of the biospecimen quality analyses, in accordance to the biobank SOPs. Finally, the biospecimens were stored in positions reserved in the biobank freezers.

Table 4.2 A list of roles and tasks assigned to each of the NSB biobank personnel.

User role	Tasks
IT system administrator	Install LIMS, create a clean Baobab site, manage the system, and create users with roles
Lab manager	Conclude agreement with the client, create project and specify the biospecimen types and analyses to be done to the client's biospecimens, and review analysis reports
Lab clerk	Manage biospecimen and inventory storage, make supplier orders, and assemble kits
Lab analyst	Create analyses, analyse biospecimens, and capture results

b. Data capture

The Baobab LIMS used during the evaluation exercise, was installed in a "built-to-specification" standalone personal computer (PC) hosted at the South African National Bioinformatics Institute (SANBI). The operating system for the PC was Ubuntu 16.04, whose official support is guaranteed until April 2021. The PC station was characterised by the following features; 32GB Random Access Memory (RAM) and 2TB disk space. Access to the machine from outside the SANBI network is restricted by a firewall and, as such, only the web server is accessible.

A clean Baobab site, named NSB, was created to track all possible information that could be extracted from the above test scenario. Baobab LIMS allows for the creation of more than one site in a Baobab LIMS installation, with each site having its own users and information. This may be important to a biobank with peripheral collection sites, and instead of implementing Baobab LIMS in every collection site, the LIMS application can be installed, for example, in the biobank IT infrastructure, and for each collection site an individual Baobab

Fig. 4.2 (a) Kit assembly process. The kit contains two empty barcoded tubes aimed to contain whole blood specimens. (b) The shipment containing the kits is ready to dispatch. (c) Waybill showing the courier information and both biobank and client addresses.

LIMS site can be created. The online user manual documentation was used as a reference guide during the data entry process (https://baobab-lims.readthedocs.io/en/latest/).

The interaction with the LIMS application began with the creation of the biobank and client users by the IT administrator. Roles, permissions and login access information were assigned to each user. The lab manager added a new project with all necessary information such as project title, description, number of participants, biospecimen types and analyses to perform on the client biospecimens. Inventory management, freezer management and order supply were responsibilities of the lab clerk. The user created 'BD EDTA' vacutainer as "Product" and placed an order for product supply. The quantities received were assigned to each product in the system and using the kit assembly module, the lab clerk added two kits with the two empty tubes. Automatic barcodes were assigned to the two tubes and were recorded in the LIMS by using the barcode scanner. A shipment record containing the two kits was added by the lab clerk. Additional information such as courier, client address and shipping dates were also provided in the shipment record. An automatic email was generated notifying the client of a possible arrival of the specific shipment from the biobank. The client used the LIMS features to acknowledge, via email, receipt of the shipment, and to notify the biobank, also via email, of the end of the process of collection. At the receipt of the biospecimens by the biobank, the lab analyst added the analyses and results of the biospecimen quality checks into the LIMS. The biospecimens were assigned storage positions in a freezer box that the lab clerk had created previously. The lab manager verified the results and generated an automatic LIMS report which was then sent as an email attachment, to the client.

c. Evaluation

The NSB group evaluated the laboratory information management system from different perspectives; (1) the workflow implementation covering the life cycle of a biospecimen from

creation to complete use or destruction, (2) security and administration management, and (3) LIMS components; kit assembly, shipping, analysis request and freezer management. A simple method based on a numbering system from 1 to 5, with 1 being very bad and 5 being excellent, was used during the assessment of the Baobab LIMS performance (Table 4.3).

Table 4.3 A simple evaluation method based on a numbering system from 1 to 5, with 1 being very bad and 5 being excellent.

Value	Description
1	Very bad
2	Bad
3	Satisfactory
4	Good
5	Excellent

d. Additional evaluations

Following the NSB practical exercise, and to further demonstrate the strengths and weaknesses of the Baobab LIMS functionalities, similar evaluations were requested from, and carried out by Institut Pasteur de Tunis and Uganda biobank. Due to the fact that the Institut Pasteur de Tunis Biobank used the online documentation to install, setup and use in a production environment Baobab LIMS, the online documentation was included in the assessments. Conversely, the Uganda Biobank technicians were trained in a face-to-face training session prior to LIMS utilisation.

4.4 Results

a. NSB evaluation

Table 4.4 summarises the results of the evaluation. All aspects were evaluated with respect to the numerical indicators (Table 4.3). The results of the NSB evaluation demonstrated that the LIMS performance did not score below 4 for any aspects evaluated (Table 4.4). One issue was reported with regards to the online user manual guide. The lab manager commented that the online documentation was "very useful for an experienced LIMS user but could be daunting for a novice user". The lab clerk, suggested an enhancement functionality to support the storing of photographs of kits and/or individual barcoded biospecimens as they were shipped and received back from clients, for quality assurance and quality control (QA/QC). He also raised concerns regarding the freezer management component, suggesting that it can be difficult for novice users and proposed that more details pertaining to the module should be elaborated on, in the online documentation. The lab analysts further suggested that the addition of deviation codes (such as empty tubes, low volumes, cracked tubes) to biospecimens in a QC module would be beneficial. The documentation has since been updated, and ongoing work is in progress to add video tutorials for novice users. The other suggestions have been created as issues in the github repository (https://github.com/BaobabLims/baobab.lims) and will be addressed in subsequent versions. Some screenshots from the LIMS user interface have been provided in the Appendix B.

b. Institut Pasteur de Tunis and Uganda biobank evaluations

Tables 4.5 and 4.6 indicate the evaluation results from Institut Pasteur de Tunis and the Uganda Biobank, respectively. The Institut Pasteur de Tunis biobank scored the online documentation a '4' and suggested to provide more detail on certain steps within the documentation. Two issues were raised regarding Kit assembly related to encoding problems (Use of French special characters) and adding kits without registering a study project. Another issue pertaining to the Freezer management module was noted, in that reagents could not be added to freezer shelves. Furthermore, one issue concerning the upload of user man-

Table 4.4 Scores and issues of the evaluation of the Baobab LIMS modules by the NSB biobank users.

Component	Description	Evaluation (1-5)	Issues
Workflow implementation	Biospecimen "life cycle" coverage, from collection to destruction, and the collection of key information such as the time the biospecimens arrived in the laboratory, company delivered, time and who accessioned biospecimens, the user that performed the analysis and the time of completion of the task	4	The one problem is that the online manual can create problems for a novice user
Security and administration management	Each NSB staff member and client will have specific assigned roles and a level of permission and security	5	No issues detected
Kit assembly	This component provides the protocol needed to assemble kits that will be used to collect biospecimens in the field	5	No issues detected
Shipping	The module ensures that the correct instructions are given to send the appropriate biospecimen containers (as packaged in the kits) to the client and subsequent email notification to alert the client of incoming kits	4	We have tested this module with our laboratory client and the processes were efficient
Analysis request	At the project level, analyses are defined for biospecimens based on the requirements of the project. Results of these analyses are registered and reported to the client	5	No issues detected
Freezer management	The module describes the physical storage in the NSB: rooms (within rooms), freezers, shelves, cryoboxes, and positions/locations	4	Documentation may confuse novice user

ual documents for laboratory instruments which were added in the Baobab LIMS, was also observed. The former concern with the instrument module was not included in the evaluation process, however, the users did bring this to the attention of the Baobab LIMS team during evaluation. The shipping and analysis request modules are not a requirement for the biobank, and as such, are marked as "Not applicable". Similarly, Uganda biobank assessed the Baobab LIMS online documentation as "4" and raised an issue regarding Plone dependency problems during the installation of the LIMS in a new Ubuntu 18.04 server machine. All the remaining modules were assessed by the Uganda biobank team as "5" (Excellent). The issues raised by Institut Pasteur de Tunis and Uganda Biobank are discussed in the next section.

4.5 Discussion

The NSB evaluation exercise was beneficial for both the biobank and the LIMS developers. The biorepository group experienced an alternative method, which was automated, organised and secure, in contrary to practices such as the utilisation of spreadsheets, which create risks associated with being stored in different locations, being accessible and easily edited by users without restrictions, and generating file versioning inconsistencies and potential loss of information. For the system developers there was a level of anticipation to see all components of the LIMS tested for the first time in a small biobank. The developers followed the exercise from the beginning to the end and as such, were on-hand for the possibility and appearance of unexpected bugs. Fortunately, the abovementioned did not occur, and all the LIMS components worked efficiently. The exercise demonstrated the ability of Baobab LIMS to be used in the management practices related to biobank information. In addition, the evaluation exercise revealed the importance of user feedback in the enhancement of certain modules, and the expansion of the LIMS with new functionalities. It is noteworthy to men-

Table 4.5 Scores and issues of the evaluation of the Baobab LIMS modules by the Institut Pasteur de Tunis biobank users.

Component	Description	Evaluation (1-5)	Issues
Online documentation	The online documentation (http://baobab-lims.readthedocs.io/en/latest/) used to install, setup and use Baobab LIMS	4	Certain described steps need more details
Workflow implementation	Biospecimen "life cycle" coverage, from collection to destruction, and the collection of key information such as the time the biospecimens arrived in the laboratory, company delivered, time and who accessioned biospecimens, the user that performed the analysis and the time of completion of the task	4	No issues detected
Security and administration management	Each NSB staff member and client will have specific assigned roles and a level of permission and security	4	No issues detected
Kit assembly	This component provides the protocol needed to assemble kits that will be used to collect biospecimens in the field	3	Problem with special characters (encoding) and we cannot add kits without having a project
Shipping	The module ensures that the correct instructions are given to send the appropriate biospecimen containers (as packaged in the kits) to the client and subsequent email notification to alert the client of incoming kits		Not applicable
Analysis request	At the project level, analyses are defined for biospecimens based on the requirements of the project. Results of these analyses are registered and reported to the client		Not applicable
Freezer management	The module describes the physical storage in the NSB: rooms (within rooms), freezers, shelves, cryoboxes, and positions/locations	3	We cannot add reagents in shelf of refrigerator

Table 4.6 Scores and issues of the evaluation of the Baobab LIMS modules by the Uganda biobank users.

Component	Description	Evaluation (1-5)	Issues
Online documentation	The online documentation (http://baobab-lims.readthedocs.io/en/latest/) used to install, setup and use Baobab LIMS	4	Dependencies libxslt and libxml posed a very big problem to the Plone installation
Workflow implementation	Biospecimen "life cycle" coverage, from collection to destruction, and the collection of key information such as the time the biospecimens arrived in the laboratory, company delivered, time and who accessioned biospecimens, the user that performed the analysis and the time of completion of the task	5	No issues detected
Security and administration management	Each NSB staff member and client will have specific assigned roles and a level of permission and security	5	No issues detected
Kit assembly	This component provides the protocol needed to assemble kits that will be used to collect biospecimens in the field	5	No issues detected
Shipping	The module ensures that the correct instructions are given to send the appropriate biospecimen containers (as packaged in the kits) to the client and subsequent email notification to alert the client of incoming kits	5	No issues detected
Analysis request	At the project level, analyses are defined for biospecimens based on the requirements of the project. Results of these analyses are registered and reported to the client	5	No issues detected
Freezer management	The module describes the physical storage in the NSB: rooms (within rooms), freezers, shelves, cryoboxes, and positions/locations	5	No issues detected

tion that many functionalities are commonly used in a multitude of biobanks, and as such, an enhancement feedback from a unique biobank may benefit other biorepositories sharing the same requirements, as well as those which may expand to incorporate such functionalities as they grow and evolve.

According to Davis and Venkatesh, (2004), the high rate failure which occurs in complex software development is caused during the requirements management processes and is greatly due to lack of user input, incomplete requirements and changing requirements (Davis and Venkatesh, 2004). By including the NSB team from the early development stage, particularly with regards to the design of the Baobab LIMS modules, a successful software product which caters to the outlined biobank requirements, was created. At the end of each module development, the module functionalities were demonstrated to the end users. The demonstrations aided with increasing the user feedback whereby partially met and/or changes in the requirements were instantly tracked. Consequently, the LIMS software was delivered in the desired time, and ultimately, the test of the complete product functionalities described above, were evaluated with the best scores.

Baobab LIMS meets the needs of a biobank and in the above test case, it was demonstrated that the different modules are ideally suited to the functions of a small-scale biobank. A number of valuable improvements have been identified in this evaluation exercise and forms the basis of a new specifications for the next version release of Baobab LIMS. The suggested improvements concern the QA/QC module and the online documentation.

The QC modifications suggested by the biobank users are (1) new functionality to store photographs of assembled kits on the system at two separate times, before shipment and reception back from a client, to determine packaging damage and/or proof of arrival and acceptance into biobank as well as distribution, and (2) QC enhancement by adding deviation codes to biospecimens in case of handling issues and assigning function of process

and/or storage. As it is possible that throughout the overall process a kit or a biospecimen container may be damaged, or a biospecimen may fail initial acceptance and rejection criteria. In these cases, it is vitally important to capture this information as it will determine whether biospecimens will be accepted, rejected or flagged. This will have an effect on the final destination of the biospecimen especially for those that have been rejected and subsequently not processed and stored even though initial capture had occurred in the LIMS. In this way, pictures of kits and biospecimens taken before shipment, and following receipt from a client, allows for visual identification of the possible step which may have resulted in the operational procedure failure. Furthermore, a deviation code can be assigned to a biospecimen which precisely defines the cause of biospecimen rejection, such as, a missing biospecimen, a broken biospecimen, or a biospecimen of insufficient volume, to cite a few possibilities. The suggested improvements to QC are of value and importance to aid in the identification of error-prone operating procedures and as such, will affect decisions to enhance the procedures or to replace them with harmonized operating procedures (SOPs). Adherence to best practices and SOPs, from a national and/or international specialised organisation such as the International Society for Biological and Environmental Repositories (ISBER), may dramatically improve the quality of annotated biospecimens and the performance of biobanking services. Tools such as the ISBER online self-assessment tool helps biobanks to determine how well they follow the ISBER best practices and identify areas that need improvement. The self-assessment exercise is an important undertaking and should be performed before applying for certification and/or accreditation to International Standards Organisation (ISO) (ISO 20387 general requirement for biobanks currently under review) and certification programs such the College of American Pathologists (CAP) Biorepository Accreditation and the CTRNet Biobank Certification Program (Betsou, 2017). Certification and accreditation are regarded by external bodies (funders, researchers, ethics committees) as a proof of professionalism and ensures quality products and services by the certified and/or

accredited biobank, and may therefore increases the biorepository usability and sustainability by attracting more users and funders (Barnes et al., 2016). Additionally, compliance to standards and adherence to best practices ensures that biospecimens originating from different accredited biobanks, to be used in a collaborative research study, are comparable in quality and exhibit reduced preanalytical variability. These combinatorial factors thereby provide high assurance of accuracy, reproducibility and comparability of the research results (Betsou, 2017).

With regards to the second area of concern described by the NSB, the original online LIMS instructional documentation was "difficult and confusing for a novice user". The documentation used during the test exercise was written in a single PDF document. It is however understandable that placing all the LIMS information in one file may overwhelm readers. The updated online documentation is now written with Sphinx, (http://www.sphinx-doc.org/en/master/), a tool that makes it easier to create logical and visually appealing online documentation, in the format of an interactive website, with a possibility of integrating tutorial videos. The documentation was subsequently restructured and re-organised into 3 parts; the installation manual, the setup manual and the user manual, and in so doing, has made the documentation clearer, easier and more understandable (https://baobab-lims.readthedocs.io/en/latest). The new version of the online documentation was evaluated as "4" (Good) by both Institut Pasteur de Tunis and Uganda biobank. However, parts of the documentation remain unclear to users and needs to be explained in more detail. For example, the users in the Institut Pasteur de Tunis biobank were not able to store reagents using the freezer management module, however, this is indeed possible in Baobab LIMS. Furthermore, including warnings of dependency conflicts in the documentation will avoid a long debugging process, which occurred in the case of the Ugandan users, who attempted to install Plone in a not-yet supported newly released Ubuntu version. Nevertheless, the improved documentation can itself be used for self-training of a user, as demonstrated by

the Institut Pasteur de Tunis biobank, who successfully implemented Baobab LIMS in a production environment using particularly the online documentation. Enhanced and proper documentation therefore decrease the need for recurrent face-to-face training sessions which requires individuals to travel to the respective sites and would therefore reduce this particular costs for the biobank. This can be of particular application for biobanks with already limited resources. The enhancement of the online documentation and the subsequent reduction in employee learning curve can therefore assist in reducing expenditures and ultimately increase user productivity and biobank sustainability (Lynch and Buckner-Hayden, 2010). However, training sessions of sufficient time periods do assist users with a better understanding of the LIMS functionalities and the applicability of these functionalities to their specific biobank workflows. The observed impact of training users was demonstrated by the scores of "5" (Excellent) from the Uganda biobank for all of the Baobab LIMS functionalities evaluated.

Other issues raised during the evaluation, in particular by the Institut Pasteur de Tunis biobank, included text encoding and document storage. The text encoding issue raised by the Institut Pasteur de Tunis biobank is related to language interpretation by Baobab LIMS. Consequently, the LIMS may not accept non-ASCII characters, such as French special characters. There is ongoing work to translate Baobab LIMS and its online documentation into French, and use unicode encoding to accept non-ASCII characters. Translation of Baobab LIMS into other languages will help non-English speaking users to decrease the learning curve and therefore quicken the use of the LIMS. Baobab LIMS provides users options for storing and linking file documents, such as user manuals for laboratory instrument records created in the LIMS. However, this functionality remains a basic one when compared to specialised document store tools. Progress in this area is the responsibility of WP2, with BIBBOX (a platform of open source applications) already integrating Baobab LIMS and tools for document storage, such as the SeedDMS (https://www.seeddms.org). Other po-

tential problems which have surfaced during the different evaluations, have been created as Github issues and will be addressed in future Baobab LIMS versions.

As previously mentioned, sustainability is the biggest challenge facing biobanks which operate under resource limitations (Abayomi et al., 2013). As such, open source LIMS may support biobanks in sustaining their operations by reducing the prohibitive cost associated with commercial LIMS in terms of license fees, implementation, support and training, and customisation (Kyobe et al., 2017). Furthermore, given the fact that practices vary depending on the type and scale of biobanks (Malm et al., 2013), customisation of LIMS may apply to other biobanking practices. The open source aspect of Baobab LIMS allows for customisation of the source code by in-house biobank IT software developers with new modules specific to their biorepositories. Furthermore, the continued development of Baobab LIMS by the principle developers creates an environment whereby a community of engaged developers can be formed and possibly work together to contribute significant enhancements, ultimately aiding in the value and extended usability of a more universal LIMS.

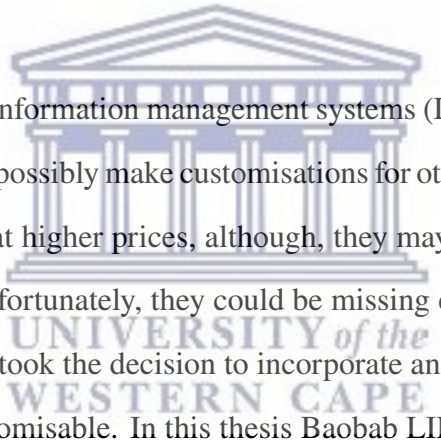
4.6 Conclusion

Evaluation exercises of a final software product by the customer (laboratory, or company) allows alignment of user requirements to the software functionalities, discovery of additional enhancements and new functionalities (that may be of great interest for customer), and also to improve robustness of the software functionalities. The evaluation exercise of Baobab LIMS at the NSB biobank allowed for discovery of two major enhancements that are indirectly related to the sustainability of the biobank; QC and online documentation. Following the success of the evaluation test, the LIMS was also demonstrated and tested in other biobanks in Tunisia and Uganda. Interestingly, the biobank team at the Institut Pasteur

de Tunis, was able to successfully install, test and start utilising the Baobab LIMS in a production environment by exclusively utilising the improved online documentation, thereby, demonstrating the importance of this resource in training and LIMS implementation. Other suggestions and valuable enhancements were obtained from the above mentioned testing in other biobank cases, and are discussed in the next chapter 5; Conclusions and further considerations.

Chapter 5

Conclusions and further considerations



The majority of laboratory information management systems (LIMS) is either developed for (1) in-house use, which can possibly make customisations for other biobanking scenarios difficult, (2) commercial sale at higher prices, although, they may not be efficient, or (3) open source sharing code but, unfortunately, they could be missing essential functionalities. The B3Africa consortium undertook the decision to incorporate an open source LIMS, which is robust, easy to use and customisable. In this thesis Baobab LIMS was developed following an evaluation by stakeholders, and with the aim to automate state-of-the-art biobank operations. One of the motivations for developing this tool was to address the lack of open source LIMS which integrates maximum functionalities required by modern biobanks. The existence of such a biobanking tool in the market, that is rich in functionalities, well-designed and free of license fees, is of great value for biobanks, and particularly for those which operate under limited resources in low and middle income countries, specifically on the African continent. A possible consequence of a profusion of a high quality open source LIMS may force commercial LIMS companies to review their prohibitive pricing strategies. Furthermore, competition between companies prevents monopoly of the market and may drive innova-

tive solutions at lowered costs. Therefore, it would benefit biobanks in the selection of an appropriate LIMS which meets the desired requirements, at lower cost.

By including the NSB team in the early stage development process of Baobab LIMS modules, the biobank requirements were clearly defined. Testing of the LIMS modules at different stages of the development was undertaken, and a system with valuable functionalities for human biobank automation was successfully delivered in the required time frames (see Chapter 3). The LIMS adheres and follows the ISBER guidelines and best practices for repository information management system (Campbell et al., 2018; Kyobe et al., 2017), and some of the valuable core functionalities are outlined below;

Authentication and authorisation: Baobab LIMS is implemented using Plone (https://plone.org/), a secure and robust Python framework. Plone has been utilised by reputable organisation such as FBI, NASA, Amnesty Internation and other well-known companies (http://plone.com/images/logo-companies-that-use-plone/view). Baobab LIMS inherits Plone security modules thereby allowing for the creation of groups of users, with permissions and roles assigned to each group.

Client - linked project registration: Collection of biospecimens can only be done within a specific project, defined by the biobank and the client. Client and project information are tracked in the LIMS.

Inventory management and supply orders : The system allows for the creation of products, placing orders with suppliers, and tracking product quantities.

Kit assembly: The process uses available quantities of products in stock and the LIMS tracks kits and products, as well as the quantities used.

Kit shipment: The shipments to, and /or from clients during biospecimen collection is monitored and logged in the system.

Biospecimen registration: Manual entry of new biospecimens, and the creation of batches of samples are both possible in Baobab LIMS. Furthermore, Baobab LIMS is MIABIS compliant and allows for biospecimens to be linked to anatomical collection site, disease ontology and donor information. With regards to donor information, regulations such as the GDPR which is related to privacy and participant data minimisation have been carefully considered.

Freezer management: Biospecimen storage is represented in the system as a hierarchical tree with the possibility of creating infinite levels of storage. Therefore, it is possible to find the exact position of a biospecimen in a storage system, for instance, the room, freezer, shelf, rack, box, and position in which a biospecimen is located. The "position" hierarchy is listed in a table and visualised in a 2D graphical representation, making it easier to locate the biospecimen.

Laboratory analyses: Clients can request for an analysis to be performed on a selected biospecimen if the service is offered by the biobank in question. The analysis information and the results obtained can be captured in the LIMS.

Report generation: Results from analyses can be automatically generated and sent to different stakeholders such as clients, laboratory managers, and/or third parties.

Log audit: Baobab LIMS uses ZODB database for back-end information storage. Records are stored as objects, and for each object type, a workflow is assigned. Transitions to a new state are automatically logged in the system. For example, chain-of-custody such as which user triggered the transition, and the date when the transition occurred is recorded. Therefore, inconsistencies in the database information can be audited. Corrective actions regarding these inconsistencies are then possible, either by manual editing of the affected records, or by restoring information to the last back-up of the database. Furthermore, if novice users to the system perform erroneous actions, log audit files will provide important metrics and may elude to further training requirements.

Database backup and restore: Baobab LIMS provides two command lines for database backup and restore; `bin/backup` and `bin/restore`, respectively. The backup of the database is a vitally important operation to undertake regularly, and it allows the restoration of LIMS information in cases where data is lost or corrupted.

In this thesis, we evaluated the developed Baobab LIMS modules in a practical test exercise. The evaluation demonstrated that the LIMS functionalities meet user requirements, and no errors were detected. However, testing of the system revealed new enhancements and functionalities, which the biobank users may not have originally been aware of. Interestingly, the enhancements were not discovered at the first communication of the biobank requirements or even at the separate module demonstrations, which motivates for testing of the combined LIMS functionalities in one comprehensive exercise. Furthermore, the exercise demonstrated the benefits of the suggested enhancements to the biobank, particularly with regards to the LIMS online documentation. Similarly, the enhancements benefit the LIMS software by increasing usability as, the changes may be valued by other biobanking institutions (see Chapter 4). Performing similar testing with other biobanks has resulted in the identification of two new enhancements, cited below;

Application Programing Interfaces (API) discovery

APIs are HTTP methods to obtain and update data recorded in the LIMS. The methods can be run using an internet browser, a terminal command line (curl), or functions in a higher level programing language (ajax function in JQuery). The API discovery module allows for the discovery of a specific, or all possible APIs, provided by Baobab LIMS. The output of the module is a json response containing information of what URI and parameters to use in a CRUD (Create, Read, Update and Delete) operation. For each parameter, information

such as name, type, and whether the parameter is required or optional, are shown in the json response. The API discovery module was developed and released in Baobab LIMS version 1.4. With the module in place, it becomes easier and faster to find an existing API to integrate into other programing applications.

Biospecimen shipment

In addition to extra shipments of biospecimens between biobanks and clients, the new enhancement will allow for intra-departmental shipments. These shipments may not require special transfer via a courier, and as such, the transfer operation can be fulfilled by an internal agent, or "runner". The module should track the transfer of a biospecimen between departments, and the turnaround time to reach the subsequent destination. This tracking of time information is necessary for quality control related to biospecimen integrity within the bioank operation. The development of the biospecimen shipment module is in its final phase and will be released soon in the next version 1.5 of Baobab LIMS.

5.1 Future considerations

Github provides metrics which present the number of clones, and number of views of the source code by unique users. The metrics obtained thus far indicate a growing interest in downloading the LIMS from the public github repository for testing (Fig. 5.1). Upon initial use of the LIMS, many laboratories and biobanks have requested availability of support from the LIMS developers. As a result, proposals and grant applications have been investigated to ensure user support availability for the future, and the creation and engagement of a developer community for further developments and enhancements of Baobab LIMS which

take into account new upcoming functionalities and technologies. Two projects are already planned for the next development phase; android and IOS mobile application which will assist in the capture of biospecimen information from collection fields that may not have internet access. Once this data has been captured the information can be stored temporarily in mobile device memories.The information transfer into Baobab LIMS can be done via API technology when the internet connection is restored. This mobile application will help assist laboratory technicians working in the field by reducing manual entry of the information in paper form, thereby, reducing risks of human errors, and turnaround time, while increasing security of data and information.

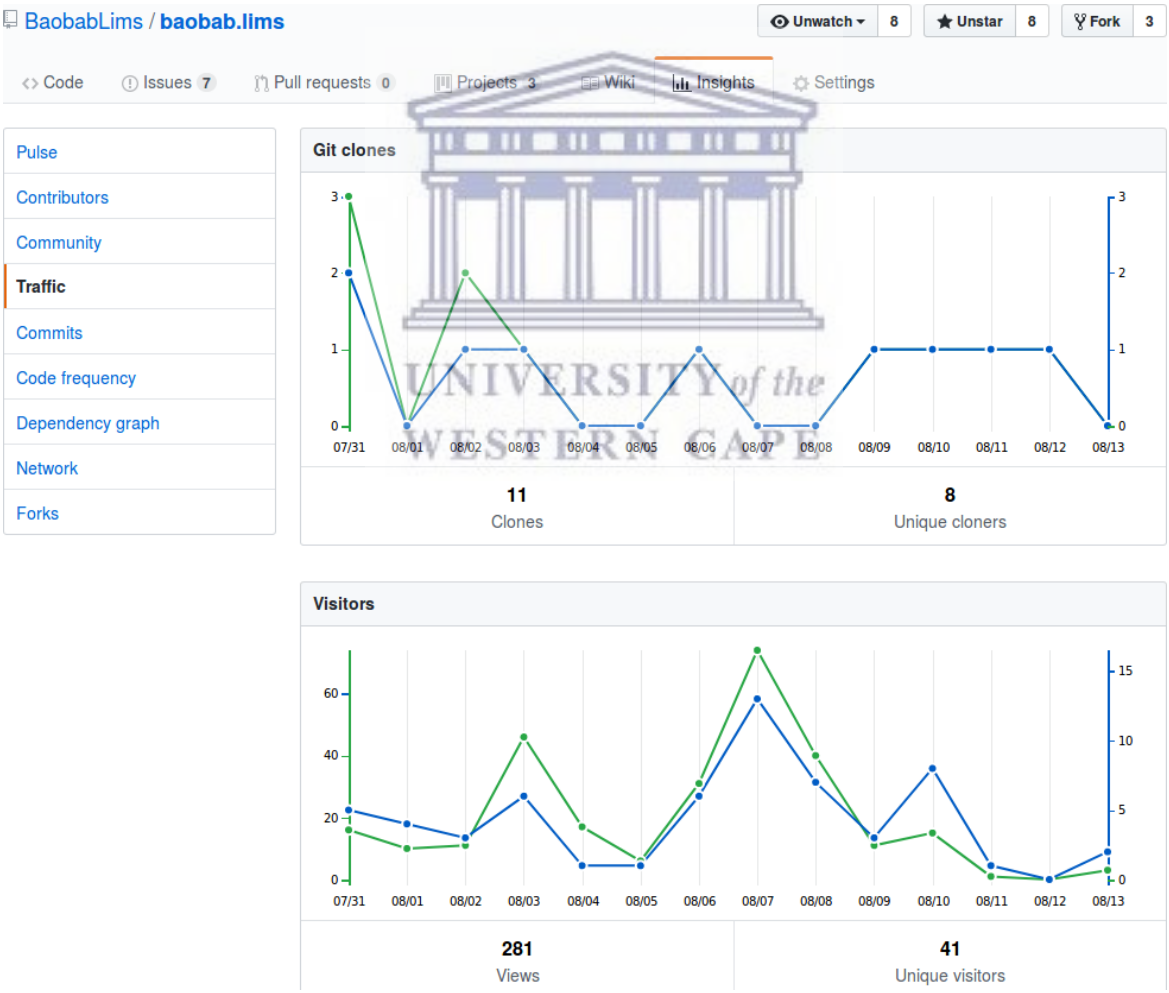


Fig. 5.1 Number of clones (downloads) and visitors, for a period of two weeks, of the Baobab LIMS source code from unique users.

The second sub-project is the workflow integration of Baobab LIMS with Stategra Experiment Management System (Hernández-de-Diego et al., 2014), and the Galaxy bioinformatics workflow management system (Afgan et al., 2016)(Fig. 5.2). It has been found that 75% of published clinical research including bioinformatics studies are not reproducible (Moore et al., 2011) and as such, this project aims to reduce possible issues related to clinical research reproducibility. Each of the abovementioned tools can be used by different groups of users; Baobab LIMS by biobank users to enter biospecimen information, Stategra EMS by laboratory technicians to enter information related to the experiments performed on the biospecimens, and Galaxy by bioinformaticians to capture information regarding the downstream bioinformatics analyses performed on the biospecimen datasets generated in the experimental phase.

In this project, a catalogue to search by biospecimen, laboratory experiment and bioinformatic analysis, will also be developed. Therefore, for a specific biospecimen used in a clinical research study, it will be possible to find the laboratory experiments performed, with the exact parameters used, and the bioinformatics workflows utilised for the analysis of the biospecimen dataset. By having access to this relevant information, reproducibility of the results may no longer be a challenging endeavour.

To this end, the future work will ultimately assist in further biobank automation, increasing productivity and quality of biospecimens. This will, consequently resulting in improved accuracy of downstream analyses with the ultimate aim of ameliorating human health and welfare.

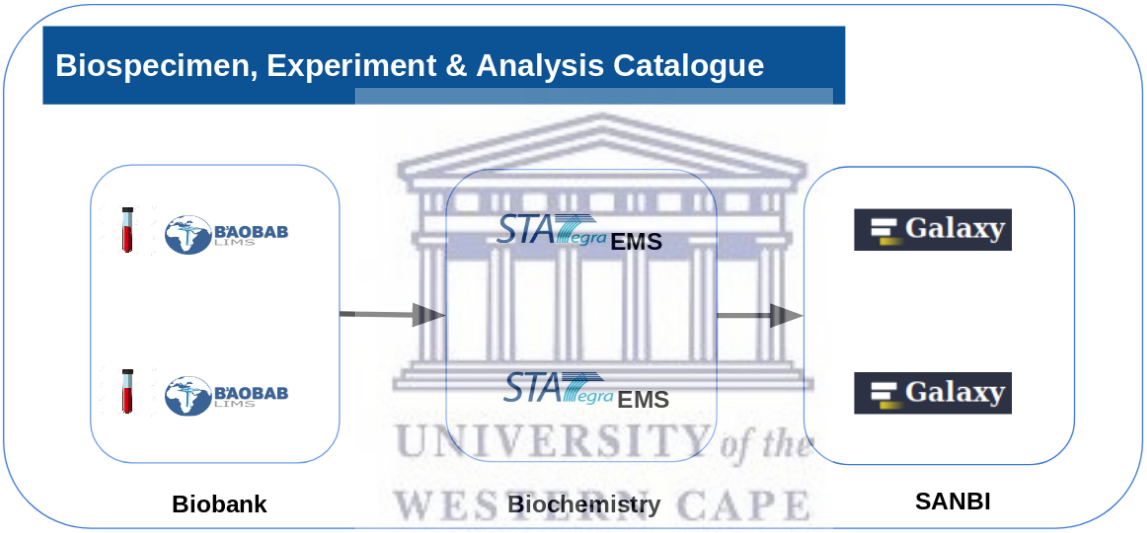


Fig. 5.2 A biobanking to bioinformatics workflow tool to track biospecimens and associated data in clinical research.

Bibliography

A. Abayomi, A. Christoffels, R. Grewal, L. A. Karam, C. Rossouw, C. Staunton, C. Swanepoel, and B. van Rooyen. Challenges of Biobanking in South Africa to Facilitate Indigenous Research in an Environment Burdened with Human Immunodeficiency Virus, Tuberculosis, and Emerging Noncommunicable Diseases. *Biopreservation and Biobanking*, 11(6):347–354, Dec. 2013. ISSN 1947-5535. doi: 10.1089/bio.2013.0049.

E. Afgan, D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Grüning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko, and J. Goecks. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, 44(W1):W3–W10, July 2016. ISSN 0305-1048. doi: 10.1093/nar/gkw343.

M. R. Anderlik. Commercial Biobanks and Genetic Research. *American Journal of Pharmacogenomics*, 3(3):203–215, June 2003. ISSN 1175-2203. doi: 10.2165/00129785-200303030-00006.

P. S. Appelbaum, C. R. Waldman, A. Fyer, R. Klitzman, E. Parens, J. Martinez, W. N. P. Ii, and W. K. Chung. Informed consent for return of incidental findings in genomic research. *Genetics in Medicine*, 16(5):367–373, May 2014. ISSN 1530-0366. doi: 10.1038/gim.2013.145.

M. O. ARSLAN. Economic Benefits of Free and Open Source Software: An Evaluation for Health Sector. page 14, 2014.

G. Avery, C. McGee, and S. Falk. Product Review: Implementing LIMS: A "how-to" guide. *Analytical Chemistry*, 72(1):57 A–62 A, Jan. 2000. ISSN 0003-2700, 1520-6882. doi: 10.1021/ac0027082.

R. O. Barnes, K. E. Shea, and P. H. Watson. The Canadian Tissue Repository Network Biobank Certification and the College of American Pathologists Biorepository Accreditation Programs: Two Strategies for Knowledge Dissemination in Biobanking. *Biopreservation and Biobanking*, 15(1):9–16, Oct. 2016. ISSN 1947-5535. doi: 10.1089/bio.2016.0021.

A. L. Barr, E. H. Young, L. Smeeth, R. Newton, J. Seeley, K. Ripullone, T. R. Hird, J. R. M. Thornton, M. J. Nyirenda, S. Kapiga, C. A. Adebamowo, A. G. Amoah, N. Wareham, C. N. Rotimi, N. S. Levitt, K. Ramaiya, B. J. Hennig, J. C. Mbanya, S. Tollman, A. A.

Motala, P. Kaleebu, and M. S. Sandhu. The need for an integrated approach for chronic disease research and care in Africa. *Global Health, Epidemiology and Genomics*, 1, 2016. ISSN 2054-4200. doi: 10.1017/gheg.2016.16.

J. G. Baust, D. Gao, and J. M. Baust. Cryopreservation. *Organogenesis*, 5(3):90–96, 2009. ISSN 1547-6278.

H. Bendou, L. Sizani, T. Reid, C. Swanepoel, T. Ademuyiwa, R. Merino-Martinez, H. Meuller, A. Abayomi, and A. Christoffels. Baobab Laboratory Information Management System: Development of an Open-Source Laboratory Information Management System for Biobanking. *Biopreservation and Biobanking*, 15(2):116–120, Apr. 2017. ISSN 1947-5543. doi: 10.1089/bio.2017.0014.

A. Berisha-Shaqiri. Management Information System And Decision-Making. *International Journal of Interdisciplinary Research SIPARUNTON*, 1(4), Oct. 2014. ISSN 2337-0572.

R. Berman. Designing and Implementing Laboratory Automation for Improved Patient Safety. *Laboratory Medicine*, 35(1):52–57, Sept. 2015. ISSN 0007-5027, 1943-7730. doi: 10.1309/3CD5V9GVWTA4RT6L.

F. Betsou. Quality Assurance and Quality Control in Biobanking. In *Biobanking of Human Biospecimens*, pages 23–49. Springer, Cham, 2017. ISBN 978-3-319-55118-0 978-3-319-55120-3. doi: 10.1007/978-3-319-55120-3_2.

F. Betsou and M. E. Sobel. The ISBER Biorepository Proficiency Testing Program: Two Successful Years Already, and New Features to Come. *Biopreservation and Biobanking*, 11(4):255–256, Aug. 2013. ISSN 1947-5535. doi: 10.1089/bio.2013.1143.

G. Bhandari and A. Snowdon. Adoption of Open Source Software in Healthcare. In *Advances in Intelligent Decision Technologies*, Smart Innovation, Systems and Technologies, pages 365–372. Springer, Berlin, Heidelberg, 2010. ISBN 978-3-642-14615-2 978-3-642-14616-9. doi: 10.1007/978-3-642-14616-9_35.

M. J. Bledsoe. Ethical Legal and Social Issues of Biobanking: Past, Present, and Future. *Biopreservation and Biobanking*, 15(2):142–147, Apr. 2017. ISSN 1947-5543. doi: 10.1089/bio.2017.0030.

P. M. Bolshete. Reporting ethics committee approval and informed consent: Review of an Ayurvedic journal. *AYU (An international quarterly journal of research in Ayurveda)*, 36 (3):254, Jan. 2015. ISSN 0974-8520. doi: 10.4103/0974-8520.182747.

E. Bravo, A. Calzolari, P. De Castro, L. Mabile, F. Napolitani, A. M. Rossi, and A. Cambon-Thomsen. Developing a guideline to standardize the citation of bioresources in journal articles (CoBRA). *BMC Medicine*, 13, Feb. 2015. ISSN 1741-7015. doi: 10.1186/s12916-015-0266-y.

L. D. Campbell, F. Betsou, D. L. Garcia, J. G. Giri, K. E. Pitt, R. S. Pugh, K. C. Sexton, A. P. Skubitz, and S. B. Somiari. Development of the ISBER Best Practices for Repositories: Collection, Storage, Retrieval and Distribution of Biological Materials for Research. *Biopreservation and Biobanking*, 10(2):232–233, Apr. 2012. ISSN 1947-5535. doi: 10.1089/bio.2012.1025.

L. D. Campbell, J. J. Astrin, Y. DeSouza, J. Giri, A. A. Patel, M. Rawley-Payne, A. Rush, and N. Sieffert. The 2018 Revision of the ISBER Best Practices: Summary of Changes and the Editorial Team's Development Process. *Biopreservation and Biobanking*, 16(1): 3–6, Feb. 2018. ISSN 1947-5535. doi: 10.1089/bio.2018.0001.

M. Campbell-Thompson, C. Wasserfall, J. Kaddis, A. Albanese-O'Neill, T. Staeva, C. Nierras, J. Moraski, P. Rowe, R. Gianani, G. Eisenbarth, J. Crawford, D. Schatz, A. Pugliese, and M. Atkinson. Network for Pancreatic Organ Donors with Diabetes (nPOD): Developing a tissue biobank for type 1 diabetes. *Diabetes/Metabolism Research and Reviews*, 28(7):608–617, Oct. 2012. ISSN 1520-7560. doi: 10.1002/dmrr.2316.

M. Capocasa, P. Anagnostou, F. D'Abramo, G. Matteucci, V. Dominici, G. D. Bisol, and F. Rufo. Samples and data accessibility in research biobanks: An explorative survey. *PeerJ*, 4:e1613, Feb. 2016. ISSN 2167-8359. doi: 10.7717/peerj.1613.

T. Caulfield and B. Murdoch. Genes, cells, and biobanks: Yes, there's still a consent problem. *PLOS Biology*, 15(7):e2002654, July 2017. ISSN 1545-7885. doi: 10.1371/journal.pbio.2002654.

S. Cervo, J. Rovina, R. Talamini, T. Perin, V. Canzonieri, P. De Paoli, and A. Steffan. An effective multisource informed consent procedure for research and clinical practice: An observational study of patient understanding and awareness of their roles as research stakeholders in a cancer biobank. *BMC Medical Ethics*, 14:30, July 2013. ISSN 1472-6939. doi: 10.1186/1472-6939-14-30.

M. Ciaburri, M. Napolitano, and E. Bravo. Business Planning in Biobanking: How to Implement a Tool for Sustainability. *Biopreservation and Biobanking*, 15(1):46–56, Feb. 2017. ISSN 1947-5535. doi: 10.1089/bio.2016.0045.

F. D. Davis and V. Venkatesh. Toward preprototype user acceptance testing of new information systems: Implications for software project management. *IEEE Transactions on Engineering Management*, 51(1):31–46, Feb. 2004. ISSN 0018-9391. doi: 10.1109/TEM.2003.822468.

J. G. Day and G. N. Stacey. Biobanking. *Molecular Biotechnology*, 40(2):202–213, Oct. 2008. ISSN 1073-6085, 1559-0305. doi: 10.1007/s12033-008-9099-7.

L. De Cecco, V. Musella, S. Veneroni, V. Cappelletti, I. Bongarzone, M. Callari, B. Valeri, M. A. Pierotti, and M. G. Daidone. Impact of biospecimens handling on biomarker research in breast cancer. *BMC Cancer*, 9:409, Nov. 2009. ISSN 1471-2407. doi: 10.1186/1471-2407-9-409.

Y. G. De Souza and J. S. Greenspan. Biobanking Past, Present and Future: Responsibilities and Benefits. *AIDS (London, England)*, 27(3):303–312, Jan. 2013. ISSN 0269-9370. doi: 10.1097/QAD.0b013e32835c1244.

J. de Vries, A. Abayomi, J. Brandful, K. Littler, E. Madden, P. Marshall, O. Ouwe Missi Oukem-Boyer, and J. Seeley. A perpetual source of DNA or something really different: Ethical issues in the creation of cell lines for African genomics research. *BMC medical ethics*, 15:60, Aug. 2014. ISSN 1472-6939. doi: 10.1186/1472-6939-15-60.

W. H. DeLone and E. R. McLean. Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research*, 3(1):60–95, Mar. 1992. ISSN 1047-7047. doi: 10.1287/isre.3.1.60.

C. Dermody, R. Irvin, R. Kadiyala, and C. Williams. Build or Buy?-The Answer Isn't so Simple Anymore. *Journal - American Water Works Association*, 98(11):62–71, Nov. 2006. ISSN 0003150X. doi: 10.1002/j.1551-8833.2006.tb07805.x.

A. Dhai. Establishing national biobanks in South Africa: The urgent need for an ethico-regulatory framework. *South African Journal of Bioethics and Law*, 6(2):38–39, Oct. 2013. ISSN 1999-7639.

R. Domack. The Future of Information Technology: Selecting a Forward-Looking LIMS Vendor. page 3.

N. Doménech García and N. Cal Purriños. Biobanks and Their Importance in the Clinical and Scientific Fields Related to Spanish Biomedical Research. *Reumatología Clínica (English Edition)*, 10(5):304–308, Sept. 2014. ISSN 2173-5743. doi: 10.1016/j.reumae.2014.02.007.

H. Dowst, B. Pew, C. Watkins, A. McOwiti, J. Barney, S. Qu, and L. B. Becnel. Acquire: An open-source comprehensive cancer biobanking system. *Bioinformatics (Oxford, England)*, 31(10):1655–1662, May 2015. ISSN 1367-4811. doi: 10.1093/bioinformatics/btv012.

S. K. Dubey, A. Anand, and H. Jangala. Laboratory information and management system: A tool to increase laboratory productivity. *Clinical Research and Regulatory Affairs*, 29 (2):46–56, June 2012. ISSN 1060-1333. doi: 10.3109/10601333.2012.692689.

M. Edward, N. Muntean, and T. Mih. LIMS use in laboratory data management. page 5, 2008.

B. S. Elger and A. L. Caplan. Consent and anonymization in research involving biobanks: Differing terms and norms present serious barriers to an international framework. *EMBO Reports*, 7(7):661–666, July 2006. ISSN 1469-221X. doi: 10.1038/sj.embor.7400740.

C. Ellervik and J. Vaught. Preanalytical variables affecting the integrity of human biospecimens in biobanking. *Clinical Chemistry*, 61(7):914–934, July 2015. ISSN 1530-8561. doi: 10.1373/clinchem.2014.228783.

L. P. Freedman, I. M. Cockburn, and T. S. Simcoe. The Economics of Reproducibility in Preclinical Research. *PLOS Biology*, 13(6):e1002165, June 2015. ISSN 1545-7885. doi: 10.1371/journal.pbio.1002165.

M. J. Green and J. R. Botkin. "Genetic exceptionalism" in medicine: Clarifying the differences between genetic and nongenetic tests. *Annals of Internal Medicine*, 138(7):571–575, Apr. 2003. ISSN 1539-3704.

M. G. Hansson. Ethics and biobanks. *British Journal of Cancer*, 100(1):8–12, Jan. 2009. ISSN 1532-1827. doi: 10.1038/sj.bjc.6604795.

H. L. Harrell and M. A. Rothstein. Biobanking Research and Privacy Laws in the United States. *The Journal of Law, Medicine & Ethics*, 44(1):106–127, Mar. 2016. ISSN 1073-1105. doi: 10.1177/1073110516644203.

K. Helsens, N. Colaert, H. Barsnes, T. Muth, K. Flikka, A. Staes, E. Timmerman, S. Wortelkamp, A. Sickmann, J. Vandekerckhove, K. Gevaert, and L. Martens. Ms_lims, a simple yet powerful open source laboratory information management system for MS-driven proteomics. *PROTEOMICS*, 10(6):1261–1264, Mar. 2010. ISSN 1615-9861. doi: 10.1002/pmic.200900409.

M. Henderson, D. Simeon-Dubach, and M. Albert. Finding the Path to Biobank Sustainability Through Sound Business Planning. *Biopreservation and Biobanking*, 13(6):385–386, Dec. 2015. ISSN 1947-5535. doi: 10.1089/bio.2015.29039.mh.

R. Hernández-de-Diego, N. Boix-Chova, D. Gómez-Cabrero, J. Tegner, I. Abugessaisa, and A. Conesa. STATegra EMS: An Experiment Management System for complex next-generation omics experiments. *BMC Systems Biology*, 8(Suppl 2):S9, Mar. 2014. ISSN 1752-0509. doi: 10.1186/1752-0509-8-S2-S9.

R. Hewitt and P. Watson. Defining biobank. *Biopreservation and Biobanking*, 11(5):309–315, Oct. 2013. ISSN 1947-5543. doi: 10.1089/bio.2013.0042.

R. E. Hewitt. Biobanking: The foundation of personalized medicine. *Current Opinion in Oncology*, 23(1):112, Jan. 2011. ISSN 1040-8746. doi: 10.1097/CCO.0b013e32834161b8.

V. Hofman, M. Ilie, E. Long, K. Washetine, C. Chabannon, D. Figarella-Branger, B. Clément, L. Mabile, A. Cambon-Thomsen, P. Boucher, G. Dagher, R. Hewitt, B. Parodi, and P. Hofman. Measuring the contribution of tumor biobanks to research in oncology: Surrogate indicators and bibliographic output. *Biopreservation and Biobanking*, 11(4): 235–244, Aug. 2013. ISSN 1947-5543. doi: 10.1089/bio.2013.0015.

C. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011):931–945, Oct. 2004. ISSN 1476-4687. doi: 10.1038/nature03001.

ISBER. 2012 best practices for repositories collection, storage, retrieval, and distribution of biological materials for research international society for biological and environmental repositories. *Biopreservation and Biobanking*, 10(2):79–161, Apr. 2012. ISSN 1947-5535. doi: 10.1089/bio.2012.1022.

G. Iskrov, T. Miteva-Katrandzhieva, and R. Stefanov. Health Technology Assessment and Appraisal of Therapies for Rare Diseases. *Advances in Experimental Medicine and Biology*, 1031:221–231, 2017. ISSN 0065-2598. doi: 10.1007/978-3-319-67144-4_13.

D. W. Jordaan. Social justice and research using human biological material: A response to Mahomed, Nöthling-Slabbert and Pepper. *SAMJ: South African Medical Journal*, 106(7): 678–680, 2016. ISSN 0256-9574. doi: 10.7196/samj.2016.v106i7.10552.

M. Joseph and G. Stone. An empirical evaluation of US bank customer perceptions of the impact of technology on service delivery in the banking sector. *International Journal of Retail & Distribution Management*, 31(4):190–202, Apr. 2003. ISSN 0959-0552. doi: 10.1108/09590550310469185.

J. Kinkorová. Biobanks in the era of personalized medicine: Objectives, challenges, and innovation. *EPMA Journal*, 7(1):4, Dec. 2016. ISSN 1878-5077, 1878-5085. doi: 10.1186/s13167-016-0053-7.

C. S. Klein. Lims User Acceptance Testing. *Quality Assurance*, 10(2):91–106, Apr. 2003. ISSN 1052-9411. doi: 10.1080/10529410390262736.

T. Klingström, M. Mendy, D. Meunier, A. Berger, J. Reichel, A. Christoffels, H. Bendou, C. Swanepoel, L. Smit, C. Mckellar-Basset, E. Bongcam-Rudloff, J. Söderberg, R. Merino-Martinez, S. Amatya, A. Kihara, S. Kemp, R. Reihs, and H. Müller. Supporting the development of biobanks in low and medium income countries. In *2016 IST-Africa Week Conference*, pages 1–10, May 2016. doi: 10.1109/ISTAFRICA.2016.7530672.

S. Kyobe, H. Musinguzi, N. Lwanga, D. Kezimbira, E. Kigozi, F. A. Katabazi, M. Wayengera, M. L. Joloba, E. A. Abayomi, C. Swanepoel, A. Abimiku, T. Croxton, P. Ozumba, A. Thankgod, A. Christoffels, L. van Zyl, E. S. Mayne, M. Kader, G. Swartz, and H3Africa Biorepository PI Working Group. Selecting a Laboratory Information Management System for Biorepositories in Low- and Middle-Income Countries: The H3Africa Experience and Lessons Learned. *Biopreservation and Biobanking*, 15(2):111–115, Apr. 2017. ISSN 1947-5535. doi: 10.1089/bio.2017.0006.

E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, Y. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church,

M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, J. Szustakowki, and International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 409(6822):860–921, Feb. 2001. ISSN 0028-0836. doi: 10.1038/35057062.

T. J. Long. Human issues and LIMS: The impact on quality. *Chemometrics and Intelligent Laboratory Systems*, 17(3):283–287, Dec. 1992. ISSN 0169-7439. doi: 10.1016/0169-7439(92)80064-B.

K. Lynch and G. Buckner-Hayden. Reducing the new employee learning curve to improve productivity. *Journal of Healthcare Risk Management*, 29(3):22–28, 2010. ISSN 2040-0861. doi: 10.1002/jhrm.20020.

B. Magzumova. Efficient Data Integrity Through LIMS-Instrument Interface. Society of Petroleum Engineers, Nov. 2016. ISBN 978-1-61399-535-8. doi: 10.2118/182564-MS.

R. R. Mahaffey. Justification for a LIMS. In *LIMS*, pages 189–204. Springer, Boston, MA, 1990. ISBN 978-1-4684-9107-4 978-1-4684-9105-0. doi: 10.1007/978-1-4684-9105-0_15.

J. Malm, T. E. Fehniger, P. Danmyr, Á. Végvári, C. Welinder, H. Lindberg, R. Appelqvist, K. Sjödin, E. Wieslander, T. Laurell, S. Hober, F. S. Berven, D. Fenyö, X. Wang, P. E. Andrén, G. Edula, E. Carlsohn, M. Fuentes, C. L. Nilsson, M. Dahlbäck, M. Rezeli, D. Erlinge, and G. Marko-Varga. Developments in biobanking workflow standardization providing sample integrity and stability. *Journal of Proteomics*, 95:38–45, Dec. 2013. ISSN 18743919. doi: 10.1016/j.jprot.2013.06.035.

P. Manders, T. M. A. Peters, A. E. Siezen, I. A. L. M. van Rooij, R. Snijder, D. W. Swinkels, and G. A. Zielhuis. A Stepwise Procedure to Define a Data Collection Framework for a Clinical Biobank. *Biopreservation and Biobanking*, 16(2):138–147, Apr. 2018. ISSN 1947-5543. doi: 10.1089/bio.2017.0084.

G. Marodin, P. França, J. C. C. da Rocha, and A. H. Campos. Biobanking for health research in Brazil: Present challenges and future directions. *Revista Panamericana de Salud Pública*, 31:523–528, June 2012. ISSN 1020-4989, 1020-4989, 1680-5348. doi: 10.1590/S1020-49892012000600012.

C. D. Mathers and D. Loncar. Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLOS Medicine*, 3(11):e442, Nov. 2006. ISSN 1549-1676. doi: 10.1371/journal.pmed.0030442.

A. Matimba, M. N. Oluka, B. U. Ebeshi, J. Sayi, O. O. Bolaji, A. N. Guantai, and C. M. Masimirembwa. Establishment of a biobank and pharmacogenetics database of African

populations. *European Journal of Human Genetics*, 16(7):780–783, July 2008. ISSN 1018-4813, 1476-5438. doi: 10.1038/ejhg.2008.49.

S. J. McCall, P. A. Branton, V. M. Blanc, S. M. Dry, J. M. Gastier-Foster, J. H. Harrison, S. D. Jewell, R. C. Dash, R. C. Obeng, J. Rose, D. L. Mateski, A. Liubinskas, J. A. Robb, N. C. Ramirez, and K. Shea. The College of American Pathologists Biorepository Accreditation Program: Results from the First 5 Years. *Biopreservation and Biobanking*, 16(1):16–22, Feb. 2018. ISSN 1947-5543. doi: 10.1089/bio.2017.0108.

R. D. McDowall. Laboratory Information Management Systems in practice. *Journal of Pharmaceutical and Biomedical Analysis*, 6(6):547–553, Jan. 1988. ISSN 0731-7085. doi: 10.1016/0731-7085(88)80068-2.

R. D. McDowall. An update on laboratory information management systems. *Journal of Pharmaceutical and Biomedical Analysis*, 11(11):1327–1330, Jan. 1993. ISSN 0731-7085. doi: 10.1016/0731-7085(93)80119-L.

L. D. McIntosh, M. K. Sharma, D. Mulvihill, S. Gupta, A. Juehne, B. George, S. B. Khot, A. Kaushal, M. A. Watson, and R. Nagarajan. caTissue Suite to OpenSpecimen: Developing an extensible, open source, web-based biobanking management system. *Journal of Biomedical Informatics*, 57:456–464, Oct. 2015. ISSN 1532-0464. doi: 10.1016/j.jbi.2015.08.020.

L. Merdad, L. Aldakhil, R. Gadi, M. Assidi, S. Y. Saddick, A. Abuzenadah, J. Vaught, A. Buhmeida, and M. H. Al-Qahtani. Assessment of knowledge about biobanking among healthcare students and their willingness to donate biospecimens. *BMC Medical Ethics*, 18, May 2017. ISSN 1472-6939. doi: 10.1186/s12910-017-0195-8.

L. Monteiro, T. Maricoto, I. S. Solha, M. Monteiro-Soares, and C. Martins. Computerised decision to reduce inappropriate medication in the elderly: A systematic review with meta-analysis protocol. *BMJ Open*, 8(1), Jan. 2018. ISSN 2044-6055. doi: 10.1136/bmjopen-2017-018988.

H. M. Moore, A. Kelly, S. D. Jewell, L. M. McShane, D. P. Clark, R. Greenspan, P. Hainaut, D. F. Hayes, P. Kim, E. Mansfield, O. Potapova, P. Riegman, Y. Rubinstein, E. Seijo, S. Somiari, P. Watson, H.-U. Weier, C. Zhu, and J. Vaught. Biospecimen Reporting for Improved Study Quality. *Biopreservation and Biobanking*, 9(1):57–70, Mar. 2011. ISSN 1947-5535. doi: 10.1089/bio.2010.0036.

N. J. Mulder, E. Adebiyi, R. Alami, A. Benkahla, J. Brandful, S. Doumbia, D. Everett, F. M. Fadlelmola, F. Gaboun, S. Gaseitsiwe, H. Ghazal, S. Hazelhurst, W. Hide, A. Ibrahimi, Y. Jaufeerally Fakim, C. V. Jongeneel, F. Joubert, S. Kassim, J. Kayondo, J. Kumuthini, S. Lyantagaye, J. Makani, A. Mansour Alzohairy, D. Masiga, A. Moussa, O. Nash, O. Ouwe Missi Oukem-Boyer, E. Owusu-Dabo, S. Panji, H. Patterton, F. Radouani, K. Sadki, F. Seghrouchni, Ö. Tastan Bishop, N. Tiffin, and N. Ulenga. H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. *Genome Research*, 26(2):271–277, Feb. 2016. ISSN 1088-9051. doi: 10.1101/gr.196295.115.

H. Müller, N. Malservet, P. Quinlan, R. Reihs, M. Penicaud, A. Chami, K. Zatloukal, and G. Dagher. From the evaluation of existing solutions to an all-inclusive package for biobanks. *Health and Technology*, 7(1):89–95, 2017. ISSN 2190-7188. doi: 10.1007/s12553-016-0175-x.

L. Norlin, M. N. Fransson, M. Eriksson, R. Merino-Martinez, M. Anderberg, S. Kurtovic, and J.-E. Litton. A Minimum Data Set for Sharing Biobank Samples, Information, and Data: MIABIS. *Biopreservation and Biobanking*, 10(4):343–348, Aug. 2012. ISSN 1947-5535. doi: 10.1089/bio.2012.0003.

J. O. Ogunbiyi, D. C. Stefan, and T. R. Rebbeck. African Organization for Research and Training in Cancer: Position and vision for cancer research on the African Continent. *Infectious Agents and Cancer*, 11(1), Dec. 2016. ISSN 1750-9378. doi: 10.1186/s13027-016-0110-9.

G. Olund, P. Lindqvist, and J. E. Litton. BIMS: An information management system for biobanking in the 21st century. *IBM Systems Journal*, 46(1):171–182, 2007. ISSN 0018-8670. doi: 10.1147/sj.461.0171.

C. Pang, F. Kelpin, D. van Enckevort, N. Eklund, K. Silander, D. Hendriksen, M. de Haan, J. Jetten, T. de Boer, B. Charbon, P. Holub, H. Hillege, and M. A. Swertz. BiobankUniverse: Automatic matchmaking between datasets for biobank data discovery and integration. *Bioinformatics*, 33(22):3627–3634, Nov. 2017. ISSN 1367-4803. doi: 10.1093/bioinformatics/btx478.

A. Park. 10 Ideas Changing the World Right Now. *Time*, 2009. ISSN 0040-781X.

B. Parodi. Biobanks: A Definition. In *Ethics, Law and Governance of Biobanking*, The International Library of Ethics, Law and Technology, pages 15–19. Springer, Dordrecht, 2015. ISBN 978-94-017-9572-2 978-94-017-9573-9. doi: 10.1007/978-94-017-9573-9_2.

C. Piggee. LIMS and the art of MS proteomics. *Analytical Chemistry*, 80(13):4801–4806, July 2008. ISSN 0003-2700, 1520-6882. doi: 10.1021/ac0861329.

P. J. Prasad and G. L. Bodhe. Trends in laboratory information management system. *Chemometrics and Intelligent Laboratory Systems*, 118:187–192, Aug. 2012. ISSN 0169-7439. doi: 10.1016/j.chemolab.2012.07.001.

P. R. Quinlan, M. Groves, L. B. Jordan, H. Stobart, C. A. Purdie, and A. M. Thompson. The Informatics Challenges Facing Biobanks: A Perspective from a United Kingdom Biobanking Network. *Biopreservation and Biobanking*, 13(5):363–370, Oct. 2015. ISSN 1947-5543. doi: 10.1089/bio.2014.0099.

W. F. Rall and G. M. Fahy. Ice-free cryopreservation of mouse embryos at -196 °C by vitrification. *Nature*, 313(6003):573–575, Feb. 1985. ISSN 1476-4687. doi: 10.1038/313573a0.

P. Ranisavljević, T. Spasić, and I. Mladenović-Ranisavljević. MANAGEMENT INFORMATION SYSTEM AND DECISION MAKING PROCESS IN ENTERPRISE. 1(3):5, 2012.

L. Rasmussen, C. B. Maddox, B. Harten, and E. L. White. A Successful LIMS Implementation: Case Study at Southern Research Institute. *JALA: Journal of the Association for Laboratory Automation*, 12(6):384–390, Dec. 2007. ISSN 1535-5535. doi: 10.1016/j.jala.2007.08.002.

M. J. Redrup, H. Igarashi, J. Schaefgen, J. Lin, L. Geisler, M. B. M'Barek, S. Ramachandran, T. Cardoso, and V. Hillewaert. Sample Management: Recommendation for Best Practices and Harmonization from the Global Bioanalysis Consortium Harmonization Team. *The AAPS Journal*, 18(2):290–293, Mar. 2016. ISSN 1550-7416. doi: 10.1208/s12248-016-9869-2.

Riegman, Morente Manuel M., Betsou Fay, de Blasio Pasquale, and Geary Peter. Biobanking for better healthcare. *Molecular Oncology*, 2(3):213–222, July 2008. ISSN 1574-7891. doi: 10.1016/j.molonc.2008.07.004.

S. Riondino, P. Ferroni, A. Spila, J. Alessandroni, R. D'Alessandro, V. Formica, D. Della-Morte, R. Palmirotta, U. Nanni, M. Roselli, and F. Guadagni. Ensuring Sample Quality for Biomarker Discovery Studies - Use of ICT Tools to Trace Biosample Life-cycle. *Cancer Genomics & Proteomics*, 12(6):291–299, 2015. ISSN 1790-6245.

Roden, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, and Masys DR. Development of a Large-Scale De-Identified DNA Biobank to Enable Personalized Medicine. *Clinical Pharmacology & Therapeutics*, 84(3):362–369, May 2008. ISSN 0009-9236. doi: 10.1038/clpt.2008.89.

S. Sarojini, A. Goy, A. Pecora, and K. S. Suh. Proactive Biobanking to Improve Research and Health Care. *Journal of Tissue Science & Engineering*, 3(2):1–6, Mar. 2012. ISSN 2157-7552. doi: 10.4172/2157-7552.1000116.

J. Scholtalbers, J. Rößler, P. Sorn, J. de Graaf, V. Boisguérin, J. Castle, and U. Sahin. Galaxy LIMS for next-generation sequencing. *Bioinformatics*, 29(9):1233–1234, May 2013. ISSN 1367-4803. doi: 10.1093/bioinformatics/btt115.

D. M. Shaw, B. S. Elger, and F. Colledge. What is a biobank? Differing definitions among biobank stakeholders. *Clinical Genetics*, 85(3):223–227, Mar. 2014. ISSN 1399-0004. doi: 10.1111/cge.12268.

K. E. Shea, E. L. Wagner, L. Marchesani, K. Meagher, and C. Giffen. Efficiently Maintaining a National Resource of Historical and Contemporary Biological Collections: The NHLBI Biorepository Model. *Biopreservation and Biobanking*, 15(1):17–19, Feb. 2017. ISSN 1947-5535. doi: 10.1089/bio.2016.0112.

M. Sheehan. Can Broad Consent be Informed Consent? *Public Health Ethics*, 4(3):226–235, Nov. 2011. ISSN 1754-9973. doi: 10.1093/phe/phr020.

D. Simeon-Dubach, A. D. Burt, and P. A. Hall. Quality really matters: The need to improve specimen quality in biomedical research. *The Journal of Pathology*, 228(4):431–433, Dec. 2012. ISSN 1096-9896. doi: 10.1002/path.4117.

Sliger. Agile project management with Scrum. https://www.pmi.org/learning/library/agile-project-management-scrum-6269, 2011.

S. Slokenberga, R. Merino Martinez, and J. Reichel. Legal and ethical governance of intercontinental biobanking : Some experiences from a H2020 project. *Förvaltningsrättslig Tidskrift*, (1):169–192, 2017.

M. E. Smith and S. Aufox. Biobanking: The Melding of Research with Clinical Care. *Current genetic medicine reports*, 1(2):122–128, June 2013. ISSN 2167-4876. doi: 10.1007/s40142-013-0014-6.

D. Strech, S. Bein, M. Brumhard, W. Eisenmenger, C. Glinicke, T. Herbst, R. Jahns, S. von Kielmansegg, G. Schmidt, J. Taupitz, and H. D. Tröger. A template for broad consent in biobank research. Results and explanation of an evidence and consensus-based development process. *European Journal of Medical Genetics*, 59(6):295–309, June 2016. ISSN 1769-7212. doi: 10.1016/j.ejmg.2016.04.002.

A. Tatem, D. Rogers, and S. Hay. Global Transport Networks and Infectious Disease Spread. *Advances in parasitology*, 62:293–343, 2006. ISSN 0065-308X. doi: 10.1016/S0065-308X(05)62009-X.

A. Troncoso. Ebola outbreak in West Africa: A neglected tropical disease. *Asian Pacific Journal of Tropical Biomedicine*, 5(4):255–259, Apr. 2015. ISSN 2221-1691. doi: 10.1016/S2221-1691(15)30340-3.

W. Ulma and D. M. Schlabach. Technical Considerations in Remote LIMS Access via the World Wide Web. *Journal of Automated Methods and Management in Chemistry*, 2005 (4):217–222, 2005. ISSN 1463-9246. doi: 10.1155/JAMMC.2005.217.

J. L. Vassy, B. R. Korf, and R. C. Green. How to know when physicians are ready for genomic medicine. *Science Translational Medicine*, 7(287):287fs19–287fs19, May 2015. ISSN 1946-6234, 1946-6242. doi: 10.1126/scitranslmed.aaa2401.

J. Vaught. Biobanking Comes of Age: The Transition to Biospecimen Science. *Annual Review of Pharmacology and Toxicology*, 56(1):211–228, Jan. 2016. ISSN 0362-1642. doi: 10.1146/annurev-pharmtox-010715-103246.

J. Vaught and N. C. Lockhart. The evolution of biobanking best practices. *Clinica Chimica Acta*, 413(19):1569–1575, Oct. 2012. ISSN 0009-8981. doi: 10.1016/j.cca.2012.04.030.

J. Vaught, J. Rogers, T. Carolin, and C. Compton. Biobankonomics: Developing a sustainable business model approach for the formation of a human tissue biobank. *Journal of the National Cancer Institute. Monographs*, 2011(42):24–31, 2011. ISSN 1745-6614. doi: 10.1093/jncimonographs/lgr009.

Á. Végvári, C. Welinder, H. Lindberg, T. E. Fehniger, and G. Marko-Varga. Biobank resources for future patient care: Developments, principles and concepts. *Journal of Clinical Bioinformatics*, 1:24, Sept. 2011. ISSN 2043-9113. doi: 10.1186/2043-9113-1-24.

T. Vora and N. Thacker. Impacts of a biobank: Bridging the gap in translational cancer medicine. *Indian Journal of Medical and Paediatric Oncology : Official Journal of Indian Society of Medical & Paediatric Oncology*, 36(1):17–23, 2015. ISSN 0971-5851. doi: 10.4103/0971-5851.151773.

S. E. Wallace. What Does Anonymization Mean? DataSHIELD and the Need for Consensus on Anonymization Terminology. *Biopreservation and Biobanking*, 14(3):224–230, May 2016. ISSN 1947-5535. doi: 10.1089/bio.2015.0119.

P. H. Watson and R. O. Barnes. A Proposed Schema for Classifying Human Research Biobanks. *Biopreservation and Biobanking*, 9(4):327–333, Oct. 2011. ISSN 1947-5535. doi: 10.1089/bio.2011.0020.

P. H. Watson, S. Y. Nussbeck, C. Carter, S. O'Donoghue, S. Cheah, L. A. Matzke, R. O. Barnes, J. Bartlett, J. Carpenter, W. E. Grizzle, R. N. Johnston, A.-M. Mes-Masson, L. Murphy, K. Sexton, L. Shepherd, D. Simeon-Dubach, N. Zeps, and B. Schacter. A Framework for Biobank Sustainability. *Biopreservation and Biobanking*, 12(1):60–68, Feb. 2014. ISSN 1947-5535. doi: 10.1089/bio.2013.0064.

T. Wilkinson, A. Ly, C. Schnier, K. Rannikmäe, K. Bush, C. Brayne, T. J. Quinn, and C. L. M. Sudlow. Identifying dementia cases with routinely collected health data: A systematic review. *Alzheimer's & Dementia*, Apr. 2018. ISSN 1552-5260. doi: 10.1016/j.jalz.2018.02.016.

H. Yu. Redefining responsible research and innovation for the advancement of biobanking and biomedical research. *Journal of Law and the Biosciences*, 3(3):611–635, Dec. 2016. ISSN 2053-9711. doi: 10.1093/jlb/lsw047.

H.-Y. E. Yu and M. L. Wilkerson. Employee Engagement Is Vital for the Successful Selection of a Total Laboratory Automation System. *Laboratory Medicine*, 48(4):e66–e74, Nov. 2017. ISSN 0007-5027. doi: 10.1093/labmed/lmx030.

Appendix A

Baobab LIMS installation

The implementation of Baobab LIMS is based on Plone framework. Therefore, dependencies needed by Baobab LIMS installation are mostly same dependencies required by Plone.

Plone Installation

Here we describe how to install Plone on Debian/Ubuntu 16.04 operating system. For an installation in a different operating system, check the Plone online documentation, https://docs.plone.org/4/en/manage/ installing/installation.html. The installation process requires users to have root privileges and basic knowledge of the Linux command lines using a terminal. If you are not familiar with UNIX operating systems, read this tutorial Linux shell https://maker.pro/linux/tutorial/basic-linux-commands-for-beginners. Please note that a single line must be completed at a time.

Plone dependencies

Plone framework requires the installation of additional system packages. Without these packages installed, Plone will not compile.

```
$ sudo apt install build-essential gcc python-dev git-core libffi-dev
$ sudo apt install libpcre3 libpcre3-dev autoconf libtool pkg-config
$ sudo apt install zlib1g-dev libssl-dev libexpat1-dev libxslt1.1
$ sudo apt install gnuplot libcairo2 libpango1.0-0 libgdk-pixbuf2.0-0
```

Download Plone Unified Installer

The Baobab LIMS is implemented and tested with Plone 4.3.11, a version released in 2016-09-12. You can download Plone 4.3.x by visiting the Plone site https://plone.org/download. Select and click on the Unified installer of your choice, or run the *wget* command line, in your terminal, with the path to the Plone version to install. For this installation we are using Plone 4.3.11;


```
$ wget --no-check-certificate https://launchpad.net/plone/4.3/4.3.11/
  +download/Plone-4.3.11-r1-UnifiedInstaller.tgz
```

If the download has been done from the Plone site, the installer would be located in the `~/Downloads` folder in the home directory. If the second option used i.e, the wget command line, the installer should be downloaded into the current directory.

Install Plone

To continue the installation, in the terminal, change directory to the folder containing the downloaded file then run the following command line to unpack the *tgz* file.

```
$ tar -xf Plone-4.3.11-r1-UnifiedInstaller.tgz
```

Change to the extracted folder in the terminal.

```
$ cd Plone-4.3.11-r1-UnifiedInstaller
```

Run the following command to install Plone

```
$ ./install.sh --target=/home/<ubuntu-user>/Plone --build-python zeo
```

where `--target` parameter is used to specify the path to the installation folder, `--build-python` will add and build Python package in your system, (this is optional if Python already installed) and finally `zeo` option will install Plone as a Client/Server application. Plone requires Python2.7 in order to operate. Run `./install.sh --help` to obtain the full list of the available parameters and their meaning.

Install Baobab LIMS

In the new folder created `/usr/local/Plone`, another folder named zeocluster can be found. This folder contains the configuration file `buildout.cfg`. Find in the configuration file the section starting with `eggs=`, and add bika.lims and baobab.lims to the existing entries.

Note:

Bika LIMS is a dependency that Baobab LIMS needs to function. Some of the modules in Baobab LIMS reference and import modules in Bika LIMS.

```
$ eggs =
    Plone
    Pillow
    bika.lims
    baobab.lims
```

Add to the section `develop=` the path to your version of Baobab LIMS and BIKA LIMS that should be already downloaded into your local machine. By convention, it is preferable to put the source code in "zeocluster/src" of your Plone installation folder.

```
$ developer =
    src/baobab.lims
    src/bika.lims
    src/graphite.theme
```

Use git clone or fork this project to have your own copy in your local machine. For developers, any change in your source code that you judge interesting and useful for the community, please create a pull request and let us know if you want to be part of the Baobab LIMS community project.

```
$ git clone https://github.com/BaobabLims/baobab.lims.git
$ git clone https://github.com/BaobabLims/bika.lims.git
$ git clone https://github.com/BaobabLims/graphite.theme.git
```


Save the file, and run `bin/buildout -n`. Buildout will download and install dependencies required by Baobab LIMS.

If you installed Plone as a root user i.e., using `sudo`, you should run the buildout command line with the user `plone_buildout`, this user is automatically created during the Plone installation;

```
sudo -u plone_buildout bin/buildout
```

Warning:

If you encounter issues of type "Packages not found", add the following line to the `[buildout]` section in the `buildout.cfg` file.

```
index = https://pypi.python.org/simple/
```

Test your installation

First, you will need to start the zeoserver, that is, the backend database server;

```
$ bin/zeoserver start
```

Thereafter, start one of the Plone clients in a debug mode, for this, run the following command;

```
$ bin/client1 fg
```

If you installed Plone as a root user, you will need to use the following commands instead;

```
$ sudo -u plone_daemon bin/zeoserver start
$ sudo -u plone_daemon bin/client1 fg
```

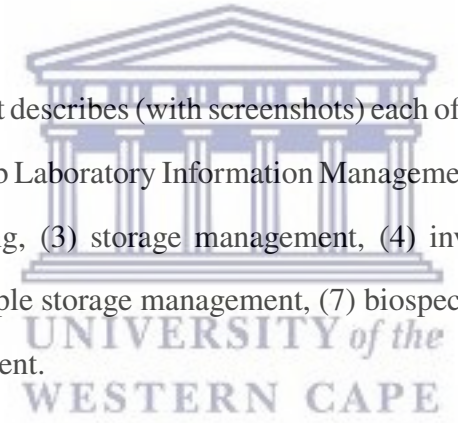
In your preferred browser, Firefox or Google chrome, run the following `http://localhost:8080` to launch Baobab LIMS application.

If installed on a remote server, an IP address (of the server) is associated with the use of LIMS eg; `https://192.168.1.1:8080`

In a production mode, other important tools need to be installed and configured, such as *Supervisorctl* and *Nginx*. The following article, https://docs.plone.org/manage/deploying/production/ubuntu_production.html details the process to follow to add and configure those tools.

Appendix B

Supplementary data



This supplementary document describes (with screenshots) each of the modules that were developed to produce the Baobab Laboratory Information Management System (LIMS) namely (1) kit assembly, (2) shipping, (3) storage management, (4) inventory management, (5) freezer management, (6) sample storage management, (7) biospecimen registration, and (8) analysis request (AR) by a client.

Kit Assembly

The following kit attributes are maintained with a kit template;

- Component list (select from components available in the database)
- Quanity of each product in a kit
- price of a kit based on the price of the kit products and VAT

The submission sites (the biobank's clients) order kits from the biobank based on a particular project to be carried out on a specific case study. For example, blood samples are to be collected to carry out DNA extraction and subsequent analysis on a group of participants. The laboratory manager navigates to "Kit assembly" form and selects from a list of prepacked/designed kit templates the template that is mostly used in the field by the client. Many variations of kit templates can be created for the DNA blood sampling kit based on the collection tubes as defined for specific downstream applications. If the appropriate kit template is not available, then the laboratory manager can create the desired kit with the appropriate collection tubes. The kit template consists of a list of components required by the client for sample collection and subsequent shipment. Supplementary Fig. B.1 shows a DNA blood sampling kit template with two components: one pair of gloves and two blood tubes.

The kit template is used to avoid the recurring selection of components during kit assembly. The kit template that consists of components is imported once for every number of kits to be assembled. In the kit assembly form (Supplementary Fig. B.2), the selected kit template will define the components and their quantities that must be added to each kit. The biobank staff member will select the specific kit template from a drop-down menu and the total number of kits that must be prepared (based on the client's requirements).

The kit assembly form specifies which stored stock items are to be used for the assembly of the kit from a list of stock items that underlies the inventory management system. The number of consumables that are required for the kit assembly is tracked and audited within the inventory management system. Biobank staff can follow this audit trail to know when consumables are running low and need to be restocked. The biobank staff members store the assembled kits in the corresponding storage under the correct conditions until the kits are shipped to the client.

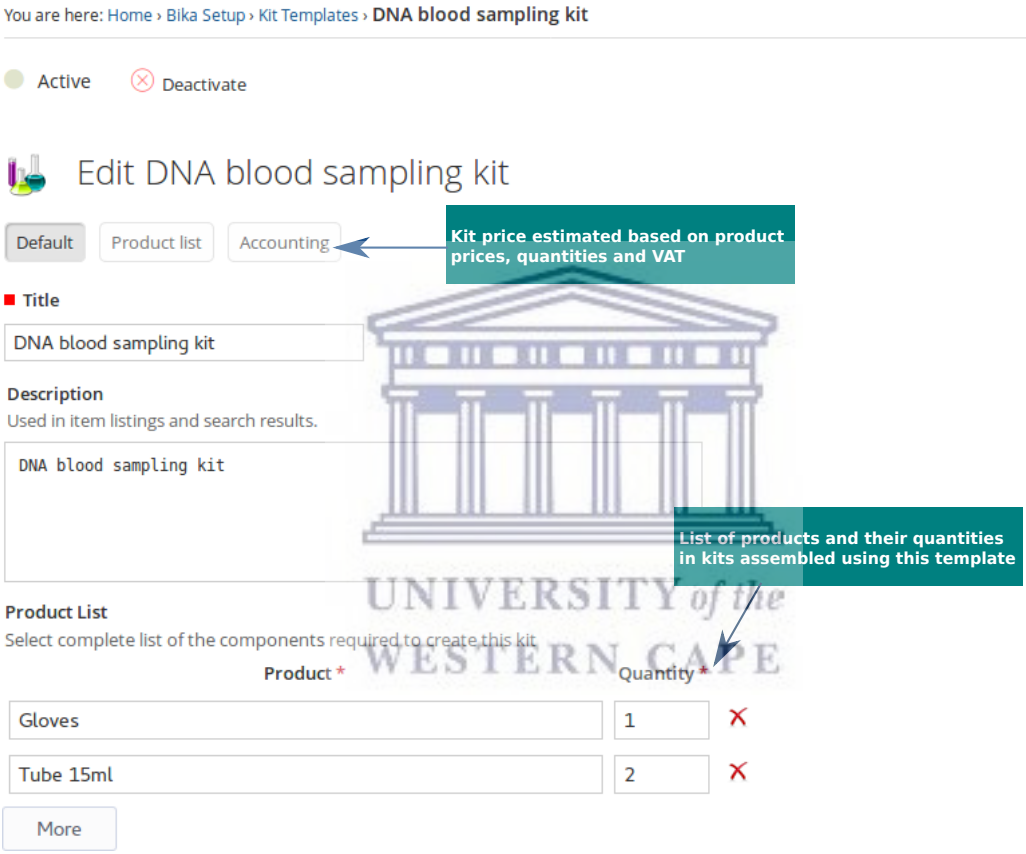


Fig. B.1 Creation of a DNA blood sampling kit. Product list field shows the list of products and their equivalent quantities to use in the kit assembly process.

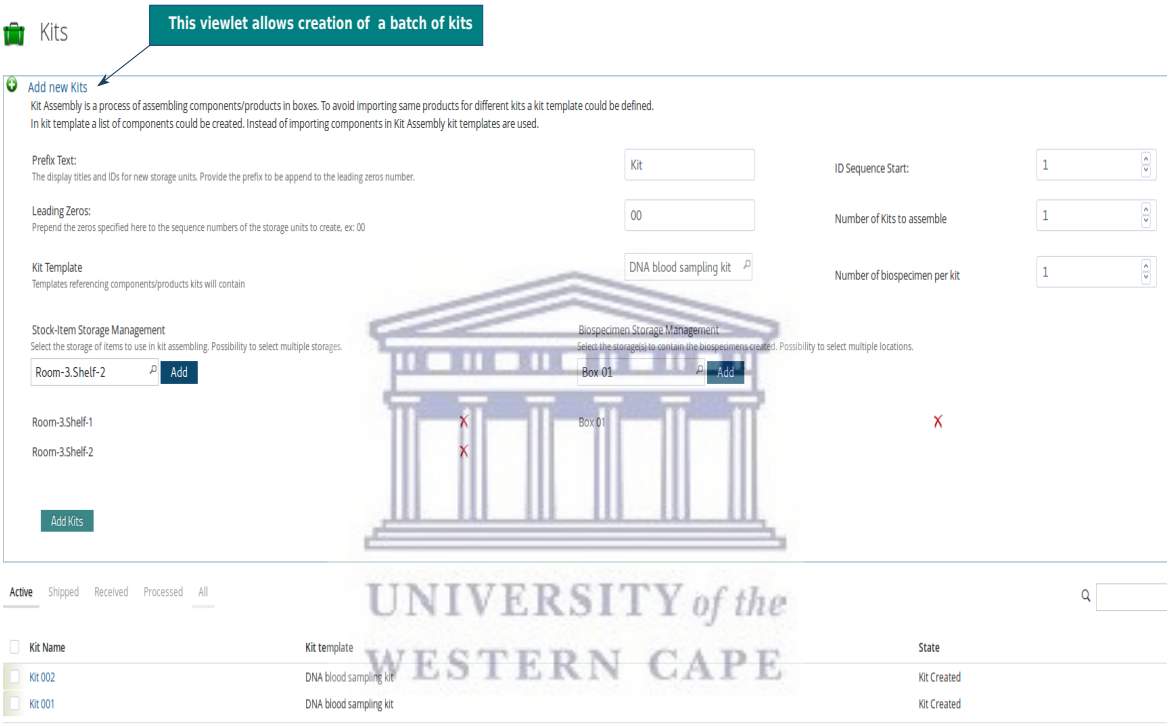


Fig. B.2 Kits creation. The specific kit template (DNA blood sampling kit) is selected and other details are added. This information is then ready for someone to physically collect materials from the storage room. Consumables are scanned to tell inventory system that a number of consumables have been released for kit assembly.

Shipping

The Biobank in consultation with the client defines the appropriate containers to ship to the client. The kits are assembled as per Kit assembly section. In our example below, one kit was prepared that contains one predesigned barcoded acid citrate dextrose collection tube with an assigned function associated with the label in a size-appropriate styrofoam box (Supplementary Fig. B.3(a)). The collection tubes are secured with laboratory tape with the barcoded label facing down. Absorbent material is placed within the cavity of the box and covers all components. A lid is added and waterproof tape used to seal the lid to the body of the box. The sealed styrofoam box is placed in a presslock plastic bag. The plastic sealed kit is placed within a corrugated shipping carton box (Supplementary Fig. B.3(b-d)) with an associated manifest (Supplementray Fig. B.4) in the pockets of the plastic bags. The courier waybill, commercial invoices, and permits if applicable are placed on the outside of courier box not covering the markings on the box. The shipping notification/manifest (Supplementary Fig. B.4) and the confirmation and query forms are sent to the biobank to notify them of an incoming shipment.

Other required regulatory documentation that accompanies all shipments may be loaded to the LIMS before shipments or is emailed to the receiving laboratory or biobank. These include the ethic approval documentation; the biospecimens deposit material transfer agreement and permits. Submission sites also have to notify the receiving laboratory and biobank of incoming shipments and prepare the following required forms before shipment. A shipment checklist is completed by the submission site and is for internal use only. The shipment manifest/notification and the shipment receipt confirmation and query form should be sent by email to the receiving site at the time of shipment. The courier's waybill number and copies of commercial invoice and permits must also be sent with the shipment.

Fig. B.3 A kit comprising collection tube and the associated components prepared for one patient or individual.

NSB
NHLS-Stellenbosch University Biobank

SHIPMENT MANIFEST

NSB Form 002
Edition 01

TO BE COMPLETED BY THE SAMPLE SUBMISSION SITE BEFORE SHIPMENT
AND RECEIPT BIOREPOSITORY AFTER RECEIPT OF SHIPMENT

SHIPMENT SENT BY		SHIPMENT RECEIVED BY	
Study Name:	[redacted]	H3A Biorepository:	[redacted]
Type of Study:	[redacted]		[redacted]
Address:	[redacted]	Address:	[redacted]
	[redacted]		[redacted]
	[redacted]		[redacted]
	[redacted]		[redacted]
Country:	[redacted]	Country:	[redacted]
Contact Person:	[redacted]	Contact Person:	[redacted]
Phone#	[redacted]	Phone#	[redacted]
E-Mail:		E-Mail:	

Number of Specimens		Batch#	Number of Specimens		Batch#
Plasma	0		Plasma		
DNA	0		DNA		
Urine	0		Urine		
Whole Blood DNA Collection Kit	1		Whole Blood DNA Collection Kit		
Shipment Conditions			**Shipment Conditions**		
Ambient	Yes		Ambient		
Frozen	N/A		Frozen		
Shipping Date	12/05/2016		Shipping Date received		
Additional Comments:			Additional Comments:		

Dry Ice "Fill" Information			
Biospecimen Submission Site *Record date, amount (kg) and time first Dry Ice Fill started*		**Host Biorepository** *Was the Dry Ice in Good condition upon receipt of shipment?*	
Date:	..... Kg Dry Ice: N/A	Yes	No (Please tick one)
Signature:		Date	
		Signature	

Fig. B.4 A shipping manifest. This form is included in the assembled kits. The client completes this form before shipping the biospecimens to the biobank.

Shipping instructions from the biobank

The "shipments" HTML page (Supplementary Fig. B.5) shows the list of shipment instructions including those that are pending (see tab in top lend hand corner). The "add" button (Supplementary Fig. B.5) pops up a new window (Supplementary Fig. B.6) with a form that captures all the instructions to ship an assembled kit to the client. The fields in this form include the details of the courier company, a Kit-ID (defined for the assembled kit) (Supplementary Fig. B.7), the date of shipping to the client, the person giving the shipping instructions, and to whom the kits will be shipped. After the request, the biobank calls the courier company to make arrangements for them to come and fetch the assembled kits.

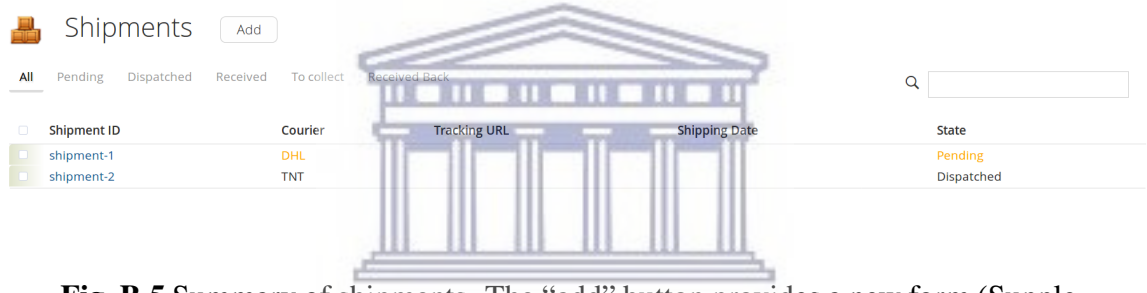


Fig. B.5 Summary of shipments. The "add" button provides a new form (Supplementary Fig. B.6) to activate the shipping instructions for an assembled kit.

Storage Management

Supplementary Fig. B.8 shows the form for storage creation. The form contains three tabs; storage units, managed storage, and unmanaged storage.

Storage units

These sections are used for creating the structure that matches the physical storage. The storage unit tab can contain multiple storage units as well as managed or unmanaged storage,

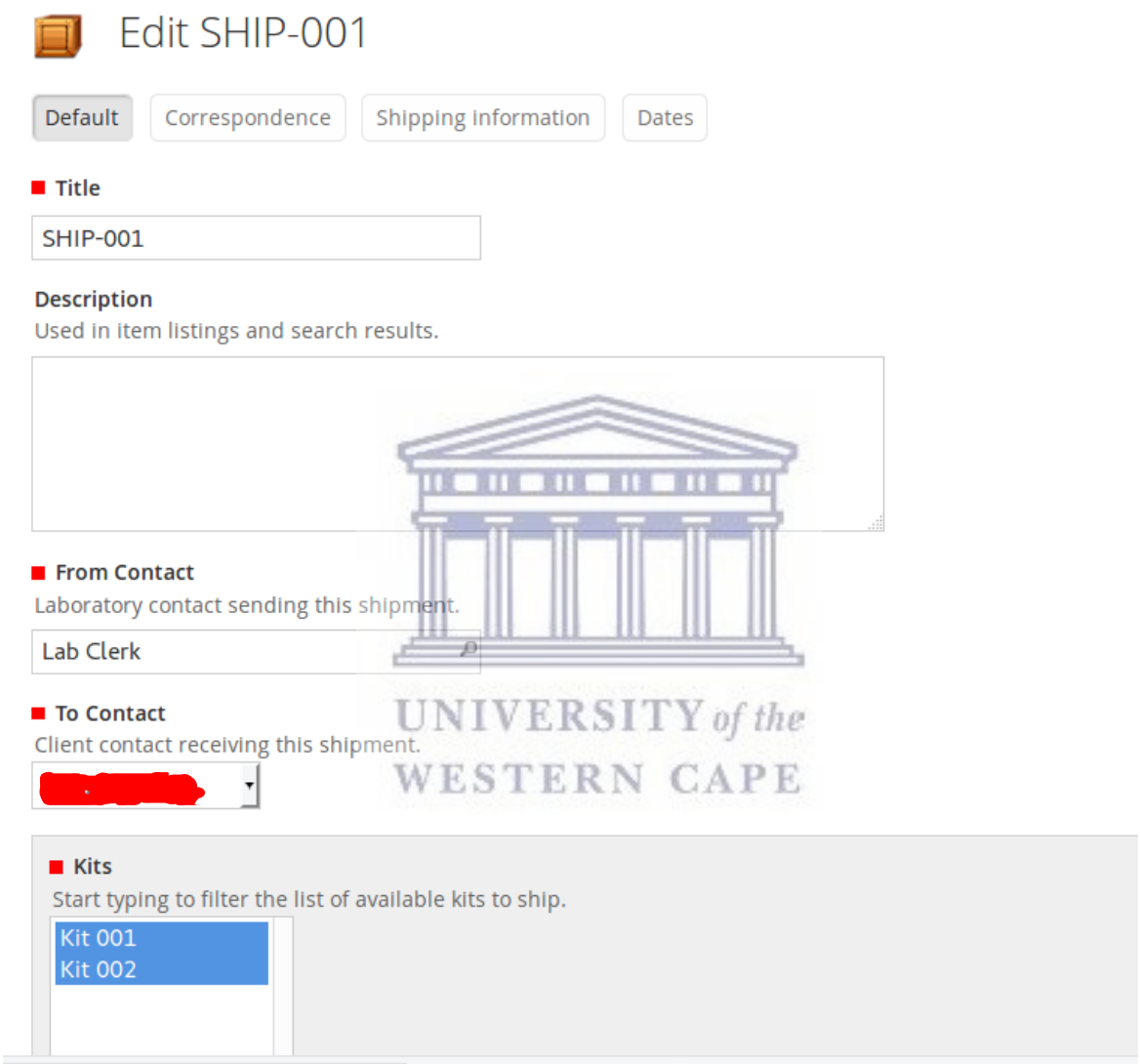
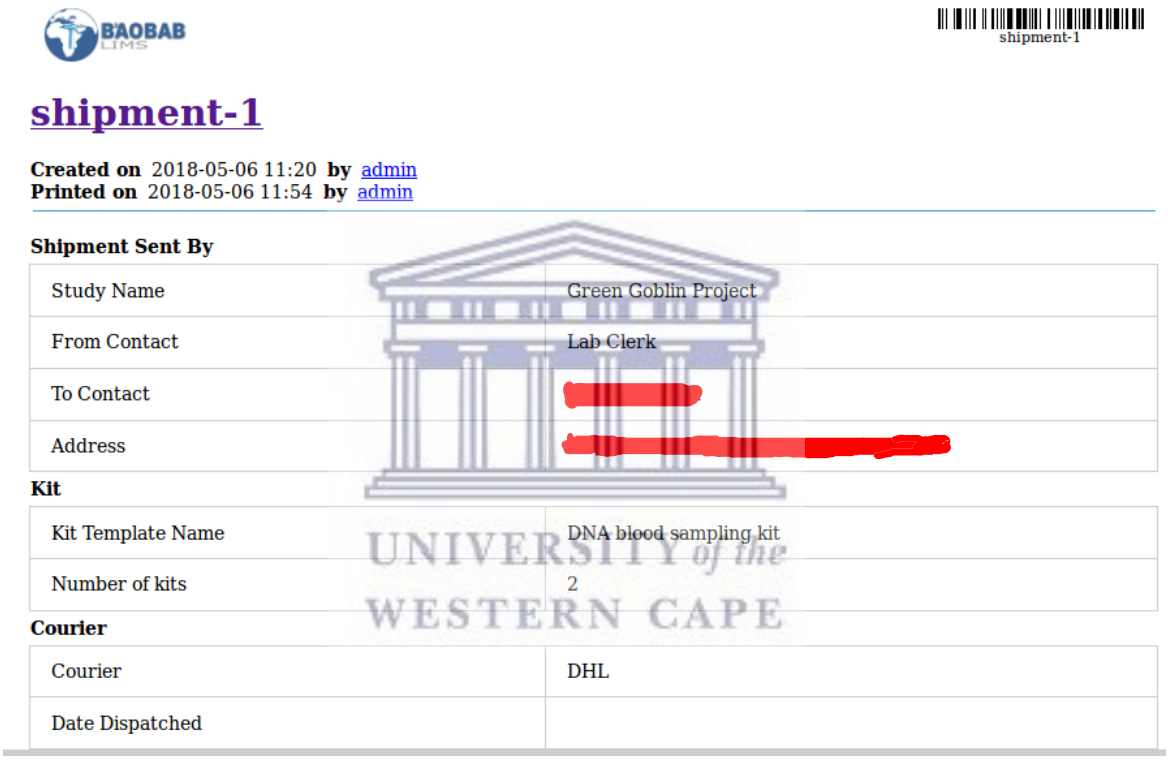


Fig. B.6 Shipping information. The assembled kits are specified in this shipping form together with the details of the courier. This form must be completed before calling the courier company.



BAOBAB LIMS

shipment-1

shipment-1

Created on 2018-05-06 11:20 **by** admin
Printed on 2018-05-06 11:54 **by** admin

Shipment Sent By

Study Name	Green Goblin Project
From Contact	Lab Clerk
To Contact	[redacted]
Address	[redacted]

Kit

Kit Template Name	DNA blood sampling kit
Number of kits	2

Courier

Courier	DHL
Date Dispatched	

Fig. B.7 Summary window for a shipping instruction. The summary shown in this figure reflects the shipping instructions as defined in Supplementary Fig. B.3. This information is for record keeping and represents the information needed for the couriers to collect the biospecimens.

but biospecimens or stock items cannot be stored directly in storage units. Storage units are defined as “room”, “freezer”, and “shelf” (Supplementary Fig. B.8).

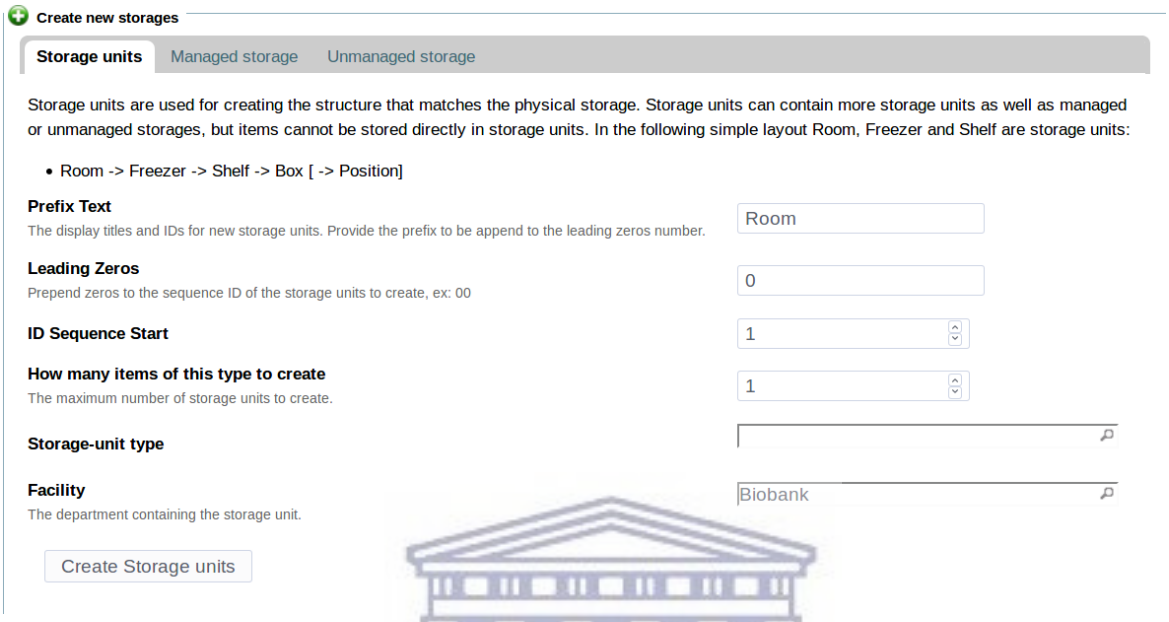


Fig. B.8 Form for creating storage units. Storage units are used for creating the structure that matches the physical storage.

Managed storage

This section contains a set number of positions for storing biospecimens, for example, boxes that can store 36 tubes each. Once all positions are occupied, the storage itself will be flagged as occupied, and when a position becomes available, the storage becomes available too. Items can be stored in specific positions, or the storage itself can be selected, in which case a position is chosen automatically—useful for storage of bulk items (Supplementary Fig. B.9).

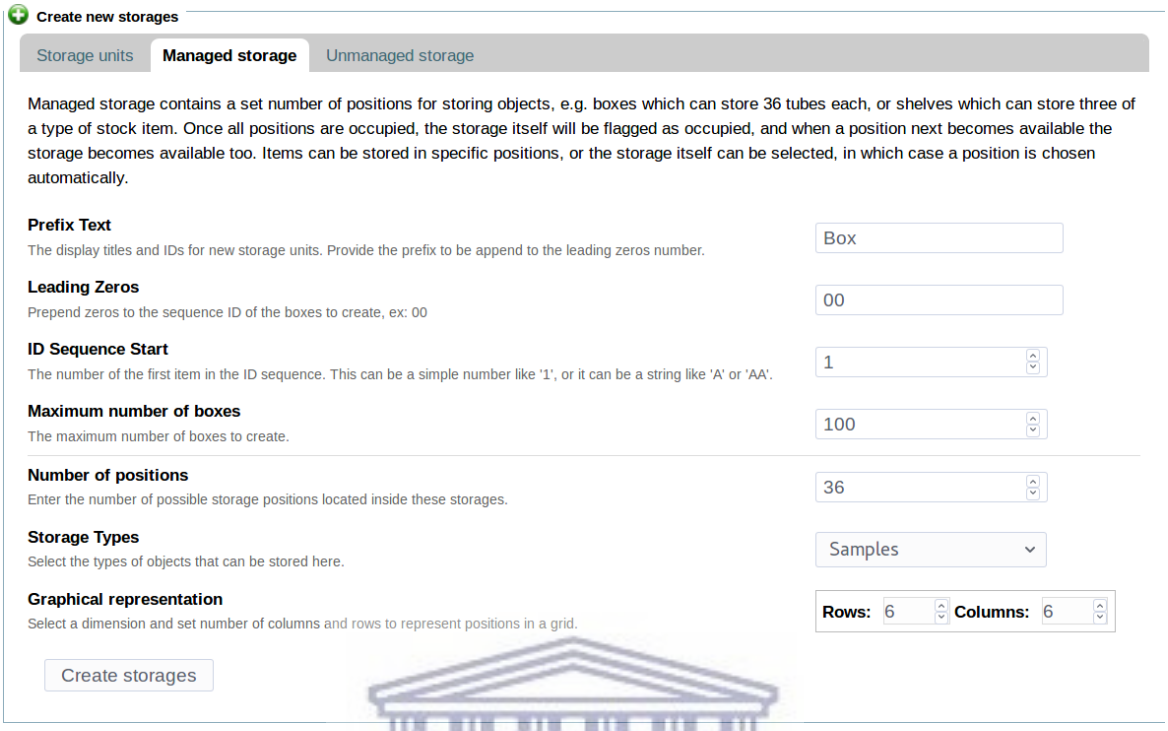


Fig. B.9 Managed storage. Box creation with fixed number of positions.

Unmanaged storage

This section does not restrict the number of items that can be stored. These storage units will be available for selection until they are manually flagged as occupied (Supplementary Fig. B.10).

Inventory storage

Stock items (products) can be stored in unmanaged storage. Unmanaged storage can be viewed as a one big location for storing stock items. There is no quanitity limit for unmanaged storage until the end user sets the location as full.

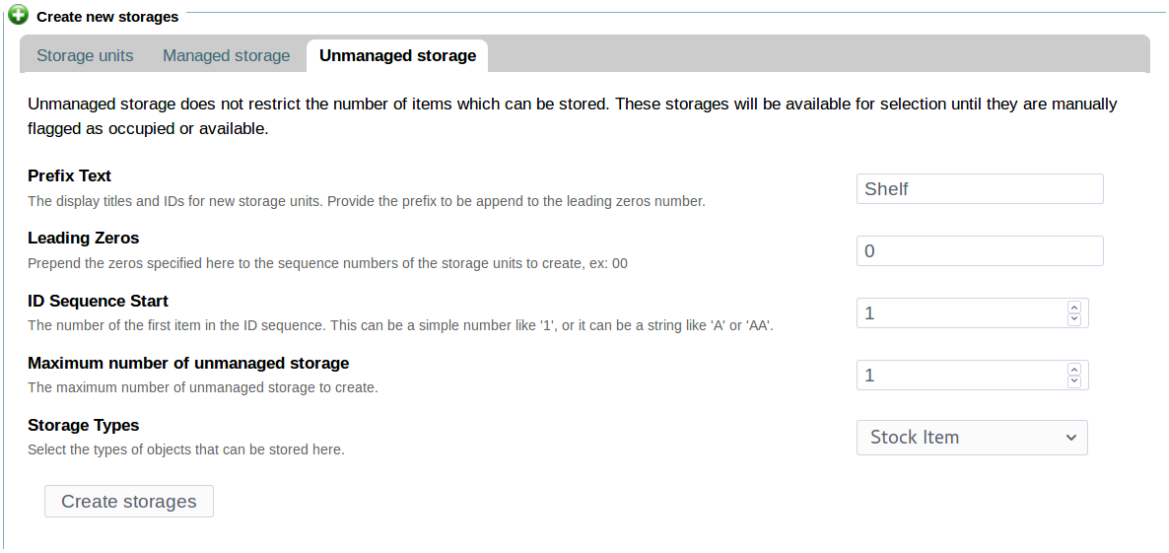


Fig. B.10 Unmanaged storage. No need for specifying the number of positions. The storage can only be manually set to "fully occupied".

Stock orders from suppliers

Stock and products should be provided before a kit is created and assembled. Specific products are ordered from a supplier. Supplementary Fig. B.11 shows the list of products that are available for a supplier called "Instruments Inc". An order is placed for 30 quantities of "blood tubes" and 5 quantities of "gloves" and depicted in Fig. B.12.

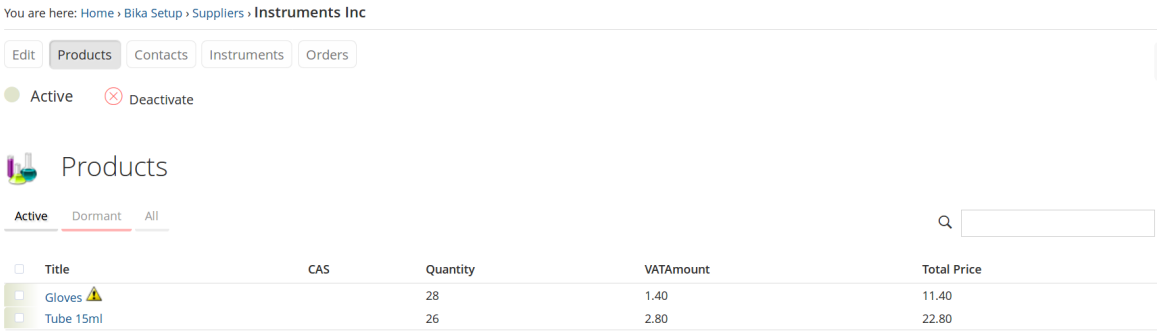


Fig. B.11 List of products available for the supplier, Instruments Inc.

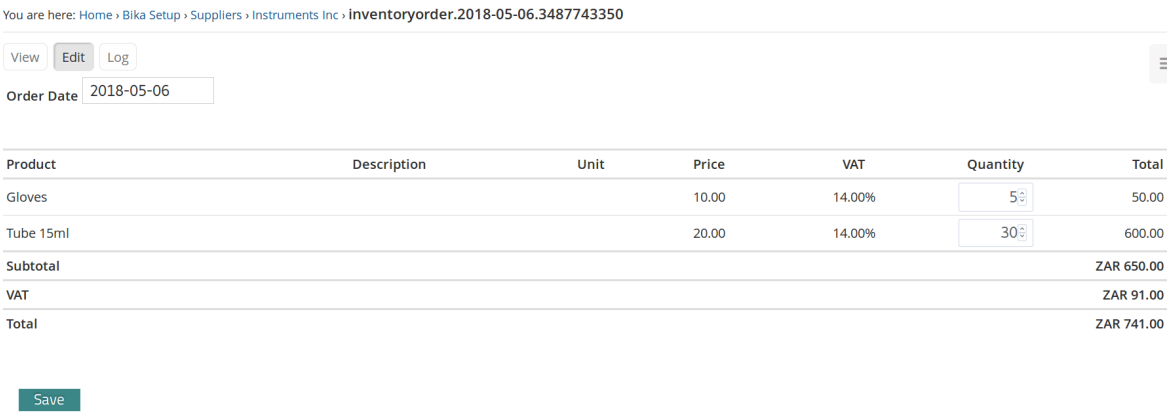


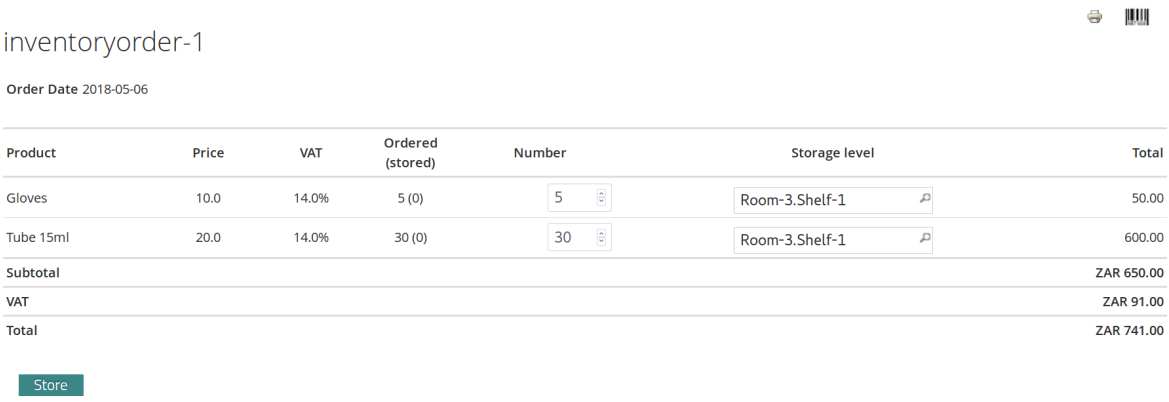
Fig. B.12 An order submitted to supplier (Instruments Inc).

Create stock items for storages

The products are automatically created as stock items after they are received from the suppliers. At this point, the stock items are ready for storage in the location defined in the next form (Fig. B.13).

Supplementary Fig. B.13 shows the precedent order when products are received. In that state, the user will be able to select the quantity and the storage location. There are scenarios where the quantities received are more than the available positions in the location selected. The system will only store the number of stock items equivalent to the number of available positions. Note that the user can select other locations if they exist. The order will be on state Stored only when all stock items are stored.

Fig. B.14 shows the content of the stock after storing the ordered product quantities.



inventoryorder-1

Order Date 2018-05-06

Product	Price	VAT	Ordered (stored)	Number	Storage level	Total
Gloves	10.0	14.0%	5 (0)	5	Room-3.Shelf-1	50.00
Tube 15ml	20.0	14.0%	30 (0)	30	Room-3.Shelf-1	600.00
Subtotal						**ZAR 650.00**
VAT						**ZAR 91.00**
Total						**ZAR 741.00**

Store

Fig. B.13 The storage levels in which stock items are stored.

Fig. B.14 Unamanaged storage after ordering. The ordered quantities stored in Room-3.Shelf-1

Freezer Management

In contrast to inventory management, freezer management follows a certain structure and order for creation. This order can be obtained only with using managed storage (Supplementary Fig. B.9).

Storage position engine

Three classes (content types) were used to design the freezer management module (Supplementary Fig. B.15) namely; Storage Unit (room), Storage Level (freezer, shelf, and box), and Storage Location (positions inside box).

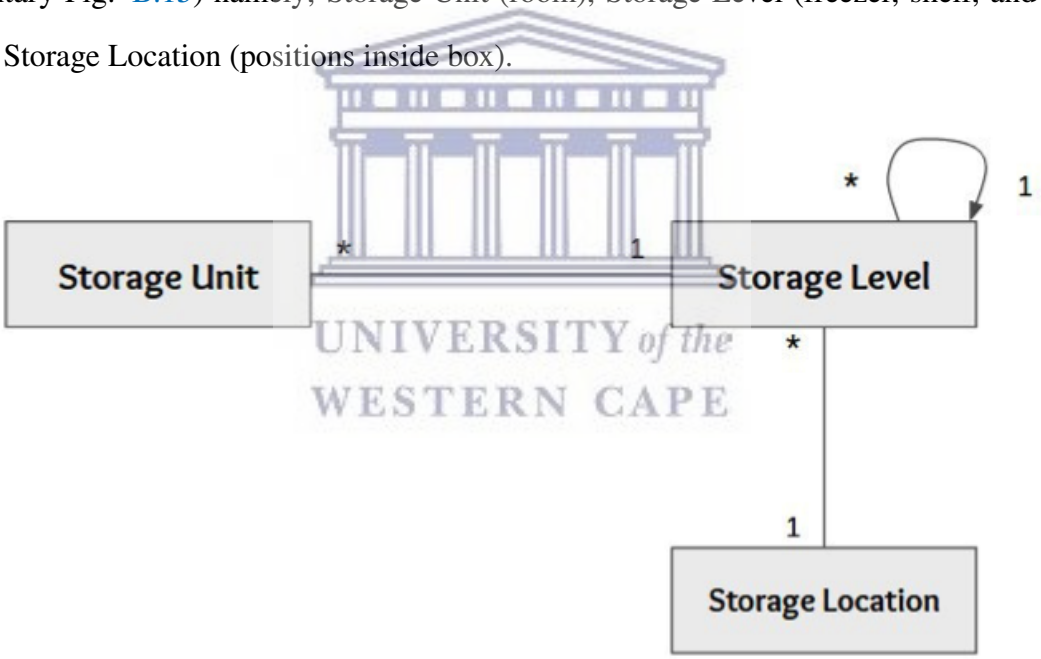


Fig. B.15 Freezer management entity relational diagram. The loop at the storage level allows the creation of infinite storage units.

Use case; Freezer configuration. Plone and Zope frameworks use ZODB, an object database for storing records (objects). Objects, by following class inheritance concept, could

be represented as a tree whereby a given object should have a parent. Supplementary Fig. B.16 shows an example for how storage is represented in a form of a hierarchical tree.

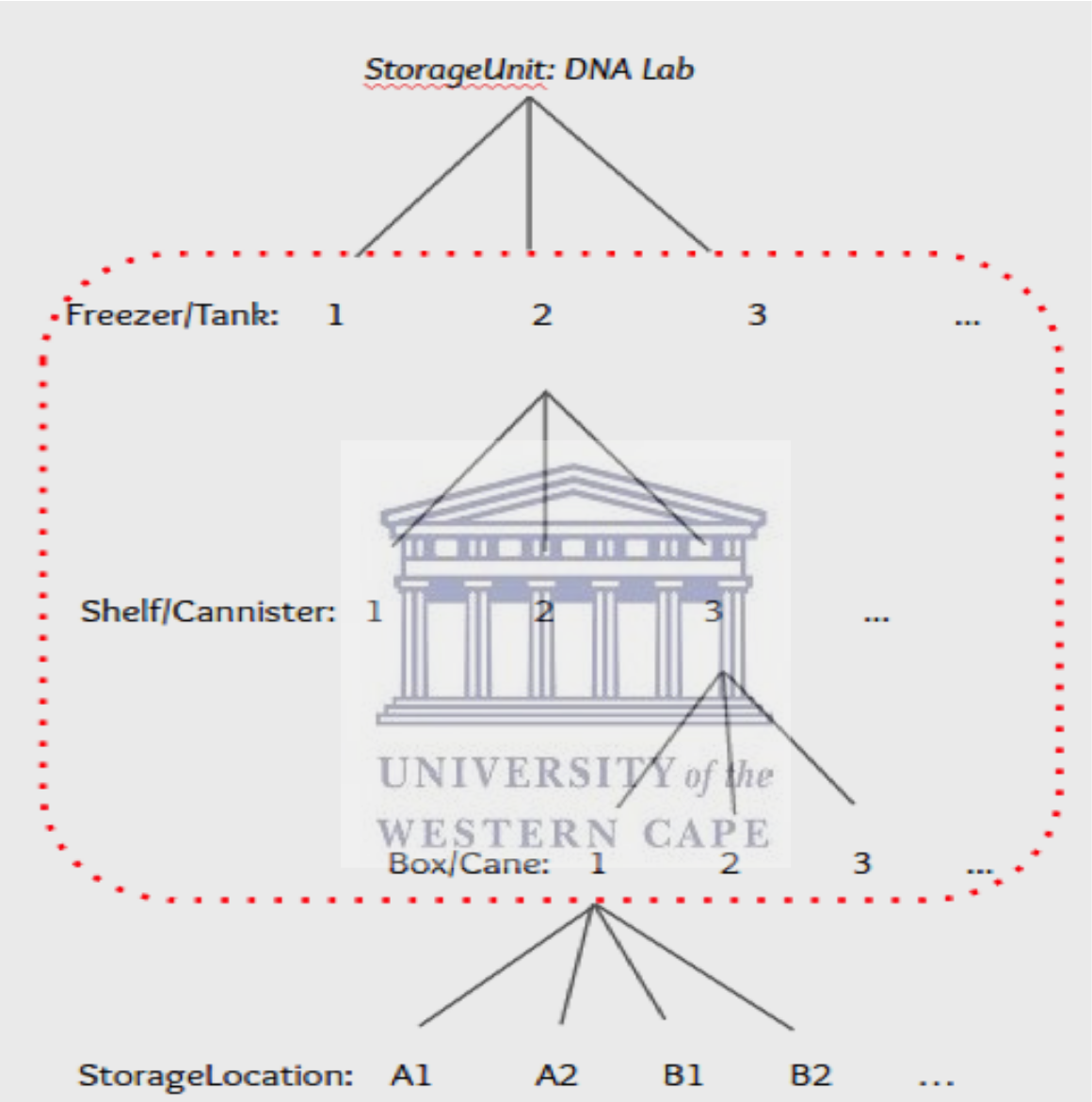


Fig. B.16 Freezer storage tree representation. The biospecimen storage is implemented in a hierarchical structure. The root of the hierarchy and the intermediate levels are storage units and the leaves are storage positions.

For both biospecimen storage and inventory management, positions are set up once for every freezer, cupboard or room, during system configuration and only again when new freezers arrive or older ones are decommissioned.

Biospecimen Storage Management

Following the structure described in the precedent section, biospecimens and aliquots can be stored in position within boxes created using “Managed storage” form (Supplementary Fig. B.9).

Graphical representation

The different storage positions for biospecimens are graphically depicted in Supplementary Fig. B.17. Each circle represents an object position. A state of a position could be “free”, “reserved”, or “occupied”. Different colors are associated to the different states; green for a free state, blue for reserved state, and red for occupied state.

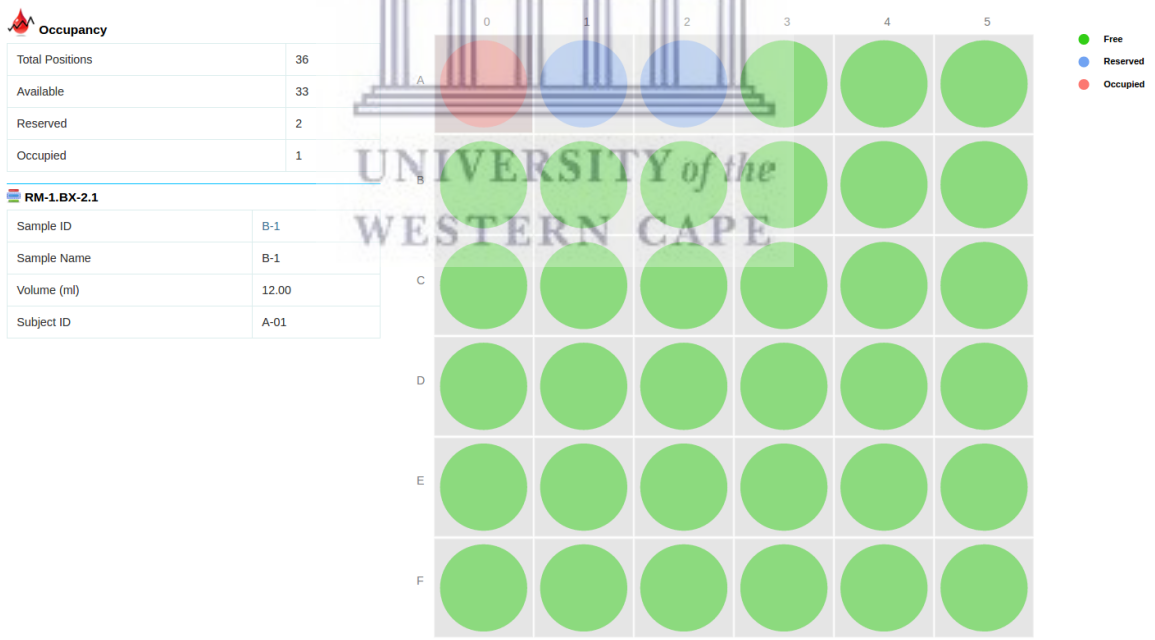


Fig. B.17 Graphical representation of a box of 36 positions. The first position is occupied and the two next are reserved for later usage.

Sample storage workflow

The following workflow was implemented to keep track of the storage position's status; First, the position created will have "Free" state. When creating a biospecimen, the selected storage position will change state to "Reserved". When the biospecimen is physically stored, the storage position state will change automatically to "Occupied" (Supplementary Fig. B.18).

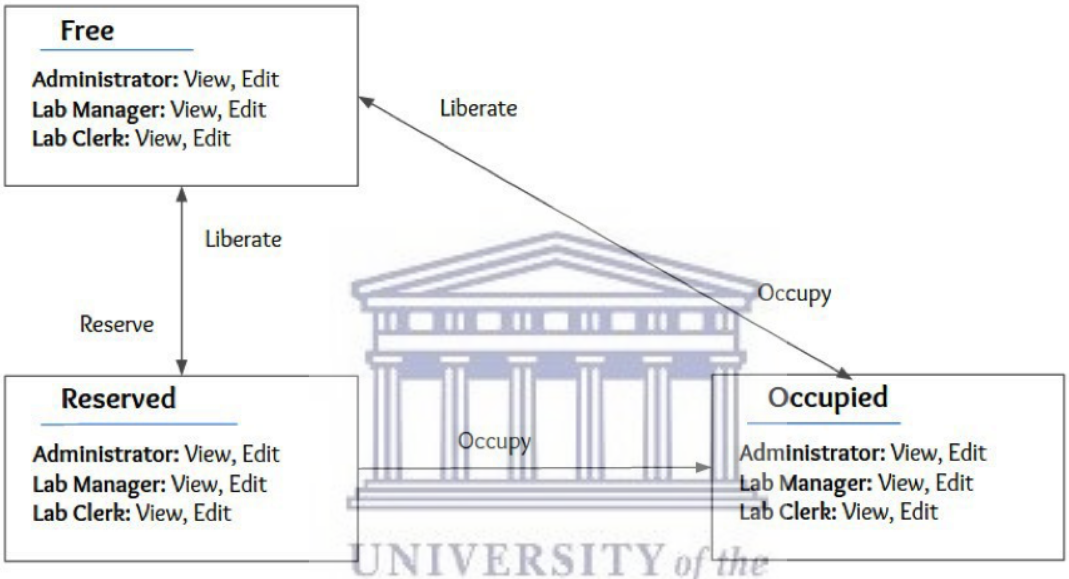


Fig. B.18 Storage location workflow. The status of a storage position is either free, reserved, or occupied. The transitions between the states can be triggered by authorised users.

Biospecimen Registration

The client returns back the kits received from the biobank after collection. A biospecimen is a material taken from human body, such as tissue, blood, plasma, stool, and urine that can be used for diagnosis and analysis. A biobank staff member opens the kits and registers the biospecimen information into the system. Baobab LIMS allows creation of unique biospec-

imen or batch of biospecimens. Supplementary Fig. B.19 is a form for creating batch of biospecimens.

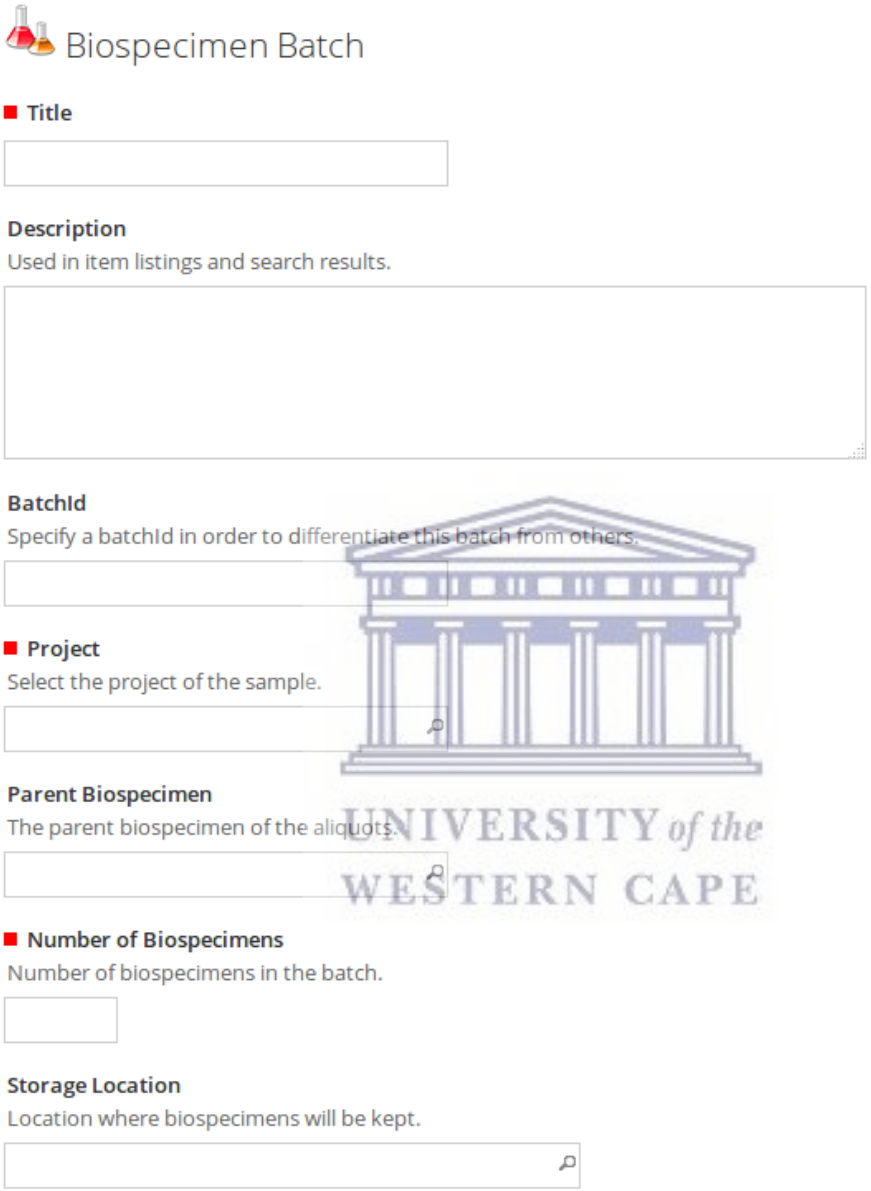


Fig. B.19 Sample batch form creation. Title, Project and the number of biospecimens are required fields.

Analysis request (AR)

The client requests for an analysis to be carried out on a specific biospecimens based on the case study of a particular project. The form used for creating AR is shown in Supplementary Fig. B.20. A biobank staff member selects the biospecimen and the analysis services for use in downstream analyses. An analyst performs the predefined experiments using laboratory instruments and the results of the analyses are therefore captured in the LIMS (Supplementary Fig. B.21).

Fig. B.20 AR form, indicating the essential fields that must be completed by the clients requesting for analyses to be carried out on a human specimen.

An instrument import interface for BioRad TC20

We implemented and integrated into Baobab LIMS a parser script for importing result files generated from BioRad TC20 instrument. The import interface allows for importing results of automated cell counting analyses that are of utility in a biomedical laboratory. A template was created for BioRad TC20 instrument interface for the analyses import form. This manages the submission of result files generated by instruments into the LIMS, which au-

tomatically imports the data after upload to avoid any form of transcription errors. BioRad TC20 instrument result files are in comma-separated value (CSV) format. The user can upload the generated files into the LIMS. By clicking on the submission button, the data in the file will be parsed and inserted in the LIMS database. The import process will significantly decrease the turnaround time and enhance accuracy of results (Supplementary Fig. B.21).

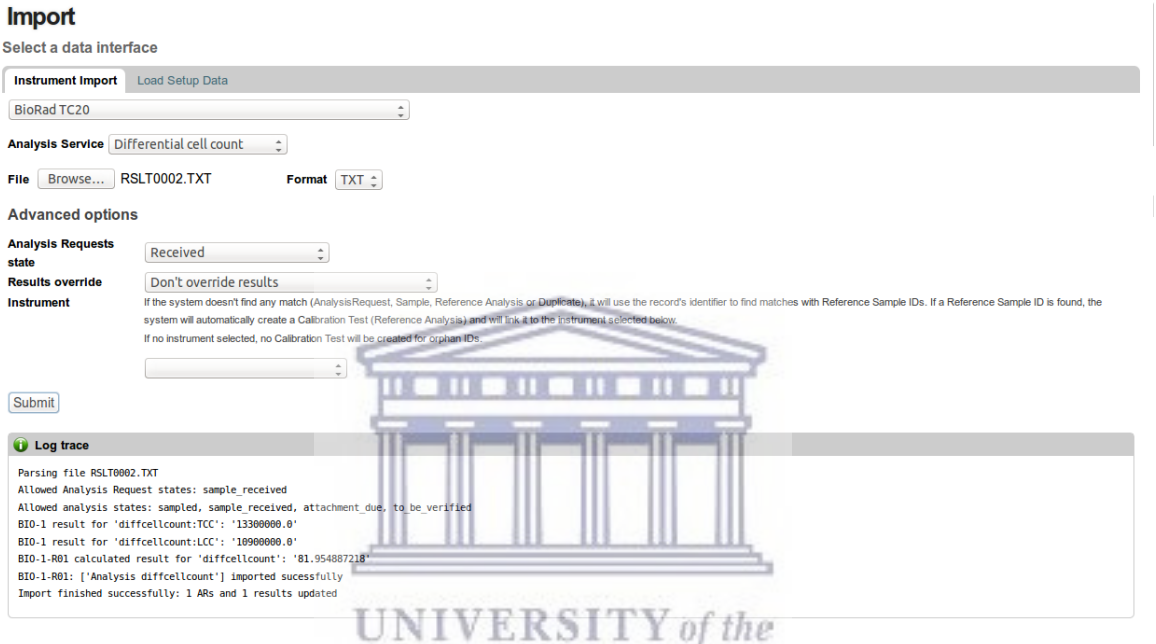


Fig. B.21 Selection of life technology instrument import interface and specifying the necessary analyses done using the instrument, and uploading the required file to be imported into the LIMS. The results are successfully imported into the LIMS database.