

Functional Requirements Overview & Phase I Objectives

version 1.1, Iemoene, 23 March 2013 **Confidential**

Table of contents

1 Purpose of this document	2
1.1 Document history	2
2 Phase I Objectives	2
2.1 Patient Demographics	3
2.2 Patient Consent	3
2.3 Kit Assembly and Distribution	3
2.4 Shipping	3
2.5 Sample Storage and Inventory	3
2.6 Analysis and Testing. LIMS	3
2.7 Document Management. SOPs	3
2.8 Accounting	3
3 Purpose of the NCB-H3A BIMS	4
3.1 BIMS Applications	4
3.2 Sustainability	4
3.3 Phased implementation	4
4 Technical Requirements	6
4.1 Open Source	6
4.2 Security. Regulatory requirements	6
4.3 Standardisation. Interoperability. Data bus	6
5 Subjects. Patients	6
6 Patient Consent. Project Ethics	7
6.1 Informed Consent	7
7 Kit Assembly and Distribution	8
8 Shipping. Distribution	8
9 Sample Storage and Inventory	8

9.1 Distribution	9
9.2 Access Policy	9
9.3 Sample Quality Management	9
9.4 Freezer Management	9
9.5 Revenue	9
9.6 Phase I	9
10 Analysis and Testing. Core LIMS	10
11 Sequencing	11
12 Stem Cells	11
13 Document Management. SOPs	11
14 Data Mining	11
15 Accounts. ERP	12
16 Glossary	13

1 Purpose of this document

This document outlines the requirements for the NHLS Cape Town H3Africa (NCB-H3A) Biobank Information Management System (BIMS) and its main components, in layman's language, maintaining scientific integrity, but for programmers and business managers to understand.

The document's purpose is to:

- Serve as background for management decisions
- Form the basis for further systems analysis towards a functional specification for the proposed system
- Demonstrate progress to the project sponsors

1.1 Document history

v 1.1 Edits after review with the BioBank team

Modified document title and added extra paragraph 2 to accentuate Phase I priorities

v 1.0 Edits shaded yellow.

After NCB team feedback and technical adviser visit

Phase I elements more clearly separated

v 0.2 Draft for NCB team feedback

v 0.1 In-house draft for feedback

2 Phase I Objectives

The Cape Town NCB-H3A Biobank project stretches over two phases. Phase I terminates in October 2014 but to satisfy project sponsors, and leave enough time to qualify for Phase II funding, the system must demonstrate 'proof of concept' and feasibility no later than 28 February 2014.

This can be achieved through the specification of all Phase I functions and their development to a point where core functionality can be demonstrated at this date.

It is expected that the NCB will start handling samples early in Phase II.

This document thus, focuses on Phase I while at the same time not excluding functionality required for Phase II, its implementation as regional biobank.

Phase I functions are discussed in context from the expanded BioBank landscape, but summarised and grouped together here:

2.1 Patient Demographics

Anonymised patient demographics as per industry standard Minimum Data Set (MDS).

2.2 Patient Consent

The BIMS stores a copy of the consent form per study only, referenced by the study's samples. All further transactions on the sample must adhere to the rules stipulated in the consent agreements.

Should a patient withdraw consent, the system must block further use of the specimen and start sample disposal and deletion of data procedures if required.

2.3 Kit Assembly and Distribution

Managing sampling kits, their assembly, components and suppliers, labels, instructions, lots and shipping.

2.4 Shipping

Sample way bill and tracking numbers hyper-linked to shipment information online. MTA and Shipping approval management. Couriers & contacts management.

2.5 Sample Storage and Inventory

The Phase I BIMS will be expected to manage: Sample Barcoding, Quality & Volume, Aliquots, Distribution, Storage location, COC,

2.6 Analysis and Testing. LIMS

Core LIMS functions:

- Genetic as well as standard clinical analysis
- Analysis batching, e.g. per research project
- Electronic laboratory notebook (ELN) functionality
- Instrument interfaces
- Quality control and diagnostics
- Reagent inventory
- Invoicing
- Document management. Lab SOP repository

2.7 Document Management. SOPs

Online document management with logs and version control.

2.8 Accounting

Invoices generated for all services, be it analyses, shipping or storage.

3 Purpose of the NCB-H3A BIMS

The sample collections stored in biobanks are invaluable and unique resources for research and clinical purposes.

The NCB BIMS has to facilitate access to biospecimens and associated data by authorized users in order to carry out genomic and proteomic analyses, while adhering to regulatory requirements. The system should also manage sample distribution, storage and shipping logistics. In many cases, the NCB won't own the samples banked, but act as 'honest broker' and storage specialists.

The BIMS must be web based and secure, to allow real-time, global access to samples and their data.

3.1 BIMS Applications

Bioinformatics is a rapidly growing field and the system has to be flexible to be used for currently defined purposes, as well as for those yet to be invented.

Currently foreseen biobank applications:

- Medical genetics.

 - Identification of genetic variations related to health and disease.

 - Drug targets and biomarkers

 - DNA sequencing for evolutionary biology including predictive and personalized medicine through disease gene identification

 - Pathogenomics

 - Therapygenetics

 - Pharmacogenetics

- Whole genome sequencing

- Population genomics studies

- DNA profiling for forensic applications

- Stem cell research, gene manipulation, targeting strategies and Bioengineering

3.2 Sustainability

It is important to keep in mind that funding for the project is currently only available up to the end of Phase II, and a further objective of the project is to ensure that it generates revenue to sustain itself thereafter.

This could be achieved not only through biobank services, including the provision of sampling kits, but also BIMS system services itself such as implementations and support.

It would not be difficult to keep the system flexible enough to also be employed for biobanks outside of the medical discipline, e.g. plant and animal samples. This will make the system much more marketable.

3.3 Phased implementation

In the bigger picture, a high volume full blooded biobank functionality will include:

- Patient demographics

- Patient consent

- Document management with version control

- Sampling kit manufacture and distribution

- Sample shipping

- Sample storage, inventory, quality and distribution

- Sample analysis, including pathological, research specific clinical and outcomes data

- Cell line management
- Sequencing
- ERP for accounts and supply chain management
- Data mining

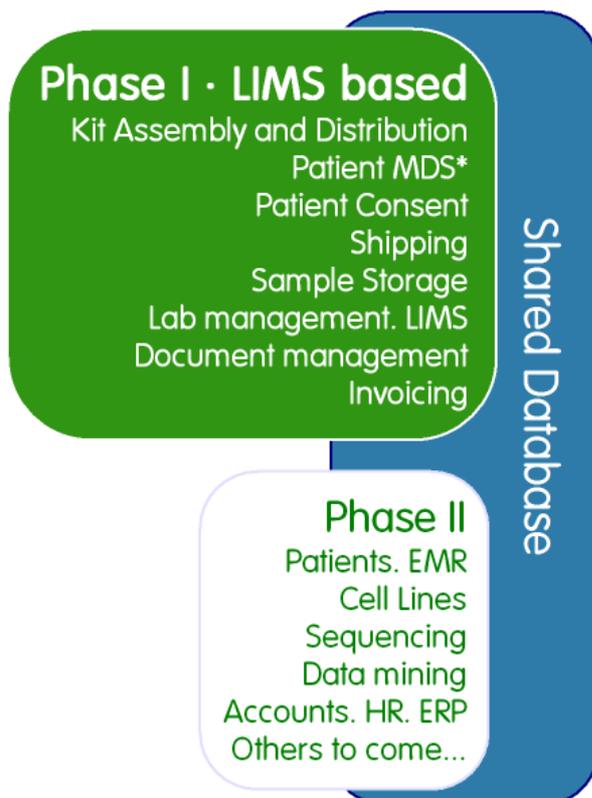
In the case of the system growing enough for these functions to be implemented as free standing but inter-connected modules, interoperability will be important and should be kept in mind from the outset and accepted interoperability standards adhered to. Please see the paragraph [Standardisation. Interoperability. Data bus.](#)

Full modularisation holds benefits for:

- Ease of use - users only have to learn the modules they most commonly use
- Performance - a modularised system can be balanced better to handle big data volumes

Given resource and time restraints, the project's Phase I objectives can in all likelihood be achieved in an expanded LIMS (laboratory information management system) without differentiating all the functions into fully fledged modules. Many currently available biobank systems, e.g. Starlims, Thermo Scientific and others, employ this 'expanded LIMS' model for their biobank systems.

Phased Implementation



*MDS - Minimum Data Set

Full blooded BIMS



4 Technical Requirements

4.1 Open Source

Though this document does not go into technical details, the NCB-H3A team wishes the proposed system to be built in Open Source technologies.

Open Source BIMS is trending, and very clear from the mountain of information openly accessible on the Internet. It stems mainly from the academic roots of many of the projects and savings and benefits to be achieved from attracting contributions from other projects without vendor lock in, closed source and license restrictions.

Robust and industry proven Linux servers provide superior security and reliability.

4.2 Security. Regulatory requirements

The biobank system will manage confidential patient data, controlled by the laws of the country where the data and samples reside, and patient privacy and data security is a high priority. Strong safeguards should be maintained to prevent unauthorized access to samples and sensitive data.

The system must comply with regulatory and audit requirements by logging data modifications and implementing appropriate authentication procedures.

There are many regulatory and best practice standards that can be applied, such as FDA 21 CFR Part 11, GMP, GCP, GTP, GxP, CAP, ISO 9000, 17025 and ISO 15189. Obtaining accreditation improves business prospects.

The browser based user interface should be secured via the strictest possible measures, e.g. digital encryption and site certification and built in security measures such as password expiry and auto log-offs applied.

The system should be platform independent, but has to support Linux for it to be secured by industry standard Unix firewalls.

4.3 Standardisation. Interoperability. Data bus

Beyond Phase I, the future BIMS could be developed as singular system incorporating all of the modules and sharing a single database. On the other end of the spectrum, a system could be built from different systems available and specialising in the functionality required by each of the BIMS modules, e.g. marrying different products for LIMS and ERP.

In the latter scenario, data will be shared between modules and it should be guarded against duplicating information in parallel data silos.

A standardised data back-bone, a data bus, might be necessary to facilitate bi-directional integration of data transfers between the different modules in a standardised way.

It is not uncommon for BIMS data to be exported to external client LIMS systems or vice versa either and interoperability standards should be adhered to regardless, and ensure integrating systems for future BIMS functions not yet foreseen, remain possible with minimum effort.

The data nomenclature and clinical terminology dictionary itself should also be standardised to an accepted system such as HL7, ICD10, Snomed CT, LOINC, etc.

5 Subjects. Patients

Many projects making use of the biobank for their studies are foreseen to do so without requiring patient demographics or medical history as they will have access to their own EMR (electronic medical records) systems. Sample in the biobank only have to be identified by their IDs.

It will not always be the case, and some researchers might want to access patient medical, immunization, drug, allergy, and family medical history as well as

their geographical, lifestyle, biographical and environmental data in the BIMS itself.

All terminology and terms associated with patients, such as diseases and signs, symptoms and social circumstances should be standardised and coded using [ICD 10](#), the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD), a medical classification list by the World Health Organization (WHO)¹.

In these cases, patient privacy must be guarded at all times and patient information anonymised and 'de-identified' as far possible.

Phase I

Anonymised Patient demographics as per industry standard MDS, 'minimum data set'.

6 Patient Consent. Project Ethics

6.1 Informed Consent

Sample and associated data should only be used for ethically and scientifically approved projects with individual consent from patients where human biological specimen are concerned. Samples should be handled per ruling regulatory requirement and in accordance with donor preference.

The goal of informed patient consent is to ensure that subjects are aware of the risks and potential benefits about participating and make a voluntary decision. He or she must be informed of the test's purpose, medical implications, alternatives, and possible risks and benefits, their privacy rights, including where their DNA will be stored and who will have access to their personal information.

An informed consent document, requiring the patient's signature, articulates all of these details. Even after signing, the patient may still opt out of the test or study; the informed consent document is not a contract and the BIMS should manage withdrawal of consent, disposal of samples, removal of data.

The Phase 1 BIMS only has to link samples to a versioned copy of the consent form that applies to it, and does not have to upload and maintain the forms signed per sample.

Consent is associated with the parent specimen as well as all downstream samples and aliquots; For both new and existing specimen for which consent was not initially obtained, e.g. routine blood samples

Phase I

The BIMS stores a copy of the consent form per study or project and samples falling under that consent agreement, reference the document on their views per hyperlink.

The consent agreement must state what the biobank is allowed to do with the samples and measures taken in the system to enforce them, e.g. if samples are not to be redistributed, the system must prevent it.

When a patient withdraws consent, the system must block further use of the specimen and start sample disposal and deletion of data procedures if required.

Before Samples can be received their consent data must be in place and signed off by the sample owners.

The BIMS should not exclude studies, e.g. on plant material, where patient consent is not required.

1 <http://apps.who.int/classifications/icd10/browse/2010/en>

7 Kit Assembly and Distribution

It is expected that the NCB will build partnerships with trusted suppliers of sample collection components and should be in a position, through bulk purchasing and using in-house expertise, to offer clients advice, quality assurance and sampling kits as service and source of revenue.

Kits can be assembled, tailored to exact client requirement, using NCB storage rooms, barcoded with expiry dates and then distributed as 'lots' and tracked using the BIMS shipping infrastructure already in place for handling samples.

Where components are received with their own barcode IDs from suppliers, full traceability can be maintained.

Phase I

Managing sampling kits, their assembly, components and suppliers, labels, instructions, lots and shipping.

8 Shipping. Distribution

The purpose of this module is to manage samples in transit between source and destination, for both receiving samples and re-distributing them to authorised recipients. Shipping is integrated with the BIMS' sample chain of custody and storage inventory.

Due to the sensitive nature of the samples, it is expected that the NCB will apply MTAs, material transfer agreements, for incoming and outgoing samples, and the BIMS must apply these business rules and reference MTAs and shipping approval in sample chain of custody.

The BIMS Shipping functionality must manage the ordering, packing and shipping of the samples and aggregate all relevant information, e.g. packing lists, courier way bills and tracking numbers etc. and temperature logs where required. It must also manage accredited couriers and their contact information.

The BIMS should capture sample shipping and handling expenditure, also internal costs, for the NCB to price these services to clients.

In many cases samples will cross borders and the BIMS shipping module must ease this process through as much automation possible, e.g. by producing documentation required for outgoing samples, and where possible interface with couriers' in-house systems.

Phase I

Sample way bill and tracking numbers hyper-linked to shipment information online.

MTA and Shipping approval management.

Maintain couriers & contacts.

9 Sample Storage and Inventory

The primary purpose of the sample storage module is to assist users to find open shelf space in the freezers to store samples and to find the stored samples again when they need to retrieve them.

Only authorised individuals should be allowed to check samples in or out and these actions be logged.

To ease storage functions, barcodes are used for both samples and storage locations. Graphical representation of storage in the system will streamline handling procedures. A search facility must be available to see which samples and empty locations are available.

The system must track sample volumes and link daughter aliquots and duplicate samples, and monitor freezing and thawing cycles, as well as disposal of them.

The BIMS tracks samples from in and out of storage to the LIMS or Shipping modules, with integrated chain of custody workflow. A specimen's full chain of custody is available at all time.

9.1 Distribution

Efficient sample distribution is said to be the most important factor for biobank success. The BIMS implements sample ordering procedures to facilitate internal and external requests for samples.

The storage module should capture a due date for the return of samples, to enable researchers to reserve specimens for a future date.

9.2 Access Policy

Ethical approval is required before samples are released for use in new projects and the BIMS must capture these approvals.

9.3 Sample Quality Management

An important function in sample storage is ensuring sample quality, e.g. validating the correct labelling of samples. It is reported that sample labelling is the biggest weak point in the sample storage chain.

If records of patient visits can be accessed, patient visits should be reconciled against the samples collected, but that is not always possible. Users working with personal patient information should not have access to test results in the BIMS.

Other quality factors that should be captured are the sample's physical condition on arrival, e.g. possible contaminations, thawing, leaks, volume, temperature, etc.

Analytical quality factors, e.g. the purity and concentration of samples, are maintained in the LIMS function of the BIMS, but should be available in sample views as quality parameters.

9.4 Freezer Management

It is expected that the NBC freezers will be monitored by their own independent management systems that'll alert staff members of power cuts and temperatures going out of range.

The BIMS' instrument management module can be used for freezer and tank maintenance and service events.

The BIMS should interface with freezers where possible to upload temperature logs, and also allow for the manual capture of temperature logs.

9.5 Revenue

Where samples are stored on behalf of NCB clients at a set fee, the BIMS should produce regular invoices for these services, say monthly, and export them to the BIMS accounts or ERP module.

9.6 Phase I

The Phase I NCB BIMS will be expected to manage

- Sample Quality & Volume

- Aliquots

- Distribution

- Storage location

- COC

10 Analysis and Testing. Core LIMS

At the centre of the BIMS, a LIMS, laboratory information management system, manages the analysis of samples and tracks them through the laboratory workflow, while also enforcing quality management and audit trails.

The LIMS facilitates genetic as well as standard clinical analyses, e.g. haematology, immunology & oncology.

Since the LIMS will mostly be used in research projects, it is important that it features electronic laboratory notebook (ELN) functionality which will allow it to capture researchers' lab notes and other relevant research material such as reference documentation and images.

The rapid advances in Bioinformatics see regular workflow changes in the lab and ideally the LIMS should allow for dynamic workflow modification by users themselves.

For genetic analyses, the BIMS LIMS must incorporate all the currently necessary workflow steps, including but not restricted to:

1. Sample Preparation

Blood sample separation/fractionation into plasma, buffy coat and red blood cells

2. Nucleic Acid Extraction

Extraction of DNA and/or RNA from sample.

3. Amplification (PCR)

Amplifying a single or a few copies of DNA or RNA for ready analysis

4. Molecular Genetic Analyses

Molecular diagnostic tests of the samples, e.g. for variations FLT3-D835, JAK2 V617F. etc.

Where during the workflow additional aliquots are created, these be recorded in the system and labelled. The system workflow must also allow for these new aliquots to be submitted to storage without necessarily being analysed further.

Some automated instruments take photos of sample passing through, helpful to determine sample volumes and origins of contamination. Where available, the LIMS uploads these too.

The LIMS uses barcoded samples and interface with lab instruments and downstream systems to eliminate human error.

Phase I

The LIMS should feature as main functions:

- Genetic as well as standard clinical analysis
- Sample inventory and storage, or interface to sample storage application. Barcoding
- Analysis batching, e.g. per research project
- Electronic laboratory notebook (ELN) functionality
- Instrument maintenance and calibration management
- Instrument interfaces
- Quality control and diagnostics
- Reagent inventory and supply chain management,
- Invoicing, interfaced with accounts or ERP applications
- Document management. Lab SOP, methods and accreditation repository

11 Sequencing

The NCB BIMS will not be expected to include sequencing or DNA profiling tools, but rather aim to integrate some of the existing platforms and their sequence data stores already available in this highly specified field.

Phase I

Hyperlinks to external systems where required.

12 Stem Cells

For stem cell applications, the BIMS is required to store cells of specific genetic lines as well as all available information related to them, e.g. track the lineage of samples and daughter and parent cells linked and the relationships graphically represented.

The BIMS' Cell Bank inherits much functionality from other BIMS modules, e.g. storage, access and chain of custody and analysis.

For the purpose of this document, it adds only management of
cell lineage data
the proliferation procedures used to multiply cells

Phase I

No requirements.

13 Document Management. SOPs

The BIMS should feature document management functionality, with version and access control, to serve as repository to all documentation related to the system, its data and samples such as SOPs, regulatory, safety and training material.

The document management module should feature flexible workflow, e.g. to create documents in private repositories, then, if required, post it for review by team members from where it can be retracted for further edits, posted for review again and finally published if that was the intention.

The system should apply access and sharing rules for private documents both per individual user and groups of users, for project groups to work and share information in private folders.

To make it easy to find information in the system, it should be able to index and search documents on keywords provided by users as well as on the content of documents in the popular formats, LibreOffice, MS Office, OpenOffice, PDF and plain text formats.

Phase I

Online document management with logs and version control.

14 Data Mining

As a singular research tool, the BIMS will only come to its full right if all of the data generated by the different modules and referenced systems can be accessed through a single user interface to set up queries, e.g. to link sequence and LIMS data to subject data for analysis.

It is a long road to that goal.

Phase I

No requirements.

15 Accounts. ERP

One of the stated objectives of the project is to be self-sustainable at the end of Phase II. For this to be realised, it is necessary to keep cost accounting in mind from the earliest design phases.

System analysts agree that invoices be generated at source where most of the information pertaining to them is available, e.g. in the BIMS for analysis services, sample storage and handling charges.

Thereafter invoices should be passed on to an accounts subsystem for further processing so as not to clutter the rest of the BIMS with accounting functionality.

This could typically be achieved by including an Enterprise Resource Planning (ERP) system as BIMS module which, apart from financials and accounts, adds further benefits such as:

- Client master data

- Inventory & Supply chain management. Suppliers

- HR. Staff qualifications. Training

Phase I

Invoices generated for all services, be it analyses, shipping or storage.

16 Glossary

For the purpose of this document:

21 CFR Part 11

21 CFR Part 11 of the Code of Federal Regulations deals with the United States Food and Drug Administration (FDA) guidelines on electronic records and electronic signatures. Part 11, as it is commonly called, defines the criteria under which electronic records and electronic signatures are considered to be trustworthy, reliable and equivalent to paper records. Wikipedia.org

BIMS - Biobank Information Management System

A biobank encompasses the entity that receives, stores, processes, and disseminates human biospecimens, their derivatives, and associated data (Wikipedia). A BIMS manages these activities in a regulated context.

CLIA - Clinical Laboratory Improvement Amendments

An United States federal regulatory standards that apply to all clinical laboratory testing performed on humans in the United States, except clinical trials and basic research. Wikipedia.org

COC - Chain of custody

Refers to the chronological documentation or paper trail, showing the seizure, custody, control, transfer, analysis, and disposition of physical or electronic evidence. Wikipedia.org

EMR - Electronic Medical Record system

Computer-based patient record and/or electronic health record library. Ahima.org

ERP - Enterprise Resource Plan

Software that integrates departments and functions across a company into one computer system, enabling various departments to share information and communicate with each other. ERP systems comprise function-specific modules designed to interact with the other modules, e.g. Accounts Receivable, Accounts Payable, Purchasing, etc. Theaccountspayablenetwork.com

FLOSS - Free/Libre/Open-Source Software

Refers to both Free Software and Open Source Software. FLOSS is liberally licensed to grant the right of users to use, study, change, and improve its design through the availability of its source code. This approach gained momentum and acceptance as the potential benefits have been increasingly recognized by both individuals and corporations. Free refers to the freedom to copy and re-use the software, rather than to the price of the software. Wikipedia.org

FLT3-ITD

The most frequent genetic alterations in acute myeloid leukaemia

FOSS - Free Open Source Software

Refers to the Free Software and Open Source communities as a whole without differentiating between the terms and the matching philosophies. Libervis.com

GTP - Good Tissue Practices

Relate to the storage and distribution of human cells, tissues and cellular and tissue-based products (HCT/Ps)

HCT/P - Human cells, tissues and cellular and tissue-based products

HL7 - Health Level 7

A specification for a health data-interchange standard designed to facilitate the transfer of health data resident on different and disparate computer systems in a health care setting. Wikipedia.org

ICD - International Statistical Classification of Diseases

A coding of diseases and signs, symptoms, abnormal findings, complaints, social circumstances and external causes of injury or diseases, as classified by the World Health Organization (WHO). Wikipedia.org.

ISO 17025

Specifies the general requirements for the competence to carry out tests and/or calibrations, including sampling. It covers testing and calibration performed using standard methods, non-standard methods, and laboratory-developed methods. ISO.org

ISO 15189 Medical laboratories

Quality management system requirements particular to medical laboratories based on ISO 17025. This working group included provision of advice to users of the laboratory service, the collection of patient samples, the interpretation of test results, acceptable turnaround times, how testing is to be provided in a medical emergency and the lab's role in the education and training of health care staff. Software solutions designed specifically for medical laboratories can aid in achieving ISO15189 certification. In particular, document control software can help by improving turnaround time, TAT, for document reviews, increasing efficiency of staff and improving overall quality. Wikipedia.org

JAK2 V617F

An acquired single nucleotide mutation in patients suffering from polycythemia vera. Ipsogen.com

LIMS · Laboratory Information Management System

Computer software that is used in the laboratory for the management of samples, laboratory users, instruments, standards and other laboratory functions such as invoicing, plate management, and work flow automation. Wikipedia.org

LOINC - Logical Observation Identifiers, Names, and Codes

A database protocol aimed at standardizing laboratory and clinical codes for use in clinical care, outcomes management, and research. Payorid.com

PCR - Polymerase Chain Reaction

A technique for amplifying DNA, making it easier to isolate, clone and sequence. Inproteomics.com

SNOMED CT - Systematized Nomenclature of Medicine -- Clinical Terms

A systematically organised computer processable collection of medical terminology covering most areas of clinical information such as diseases, findings, procedures, micro-organisms, substances, etc. It allows a consistent way to index, store, retrieve, and aggregate clinical data across specialties and sites of care. Wikipedia.org

SOP · Standard Operating Procedure

Documents that describe a specific method of accomplishing a task that is to be followed precisely the same way every time. Madison.k12

UI - User Interface

The system by which people interact with a computer - e.g. a set of commands or menus in a program. Graphical user interfaces, GUIs, use windows, icons, and pop-up menus