



Designation: E1578 – 18

Standard Guide for Laboratory Informatics¹

This standard is issued under the fixed designation E1578; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide helps describe the laboratory informatics landscape and covers issues commonly encountered at all stages in the life cycle of laboratory informatics from inception to retirement. It explains the evolution of laboratory informatics tools used in today's laboratories such as laboratory information management systems (LIMS), laboratory execution systems (LES), laboratory information systems (LIS), electronic laboratory notebooks (ELN), scientific data management systems (SDMS), and chromatography data systems (CDS). It also covers the relationship (interactions) between these tools and the external systems in a given organization. The guide discusses supporting laboratory informatics tools and a wide variety of the issues commonly encountered at different stages in the life cycle. The subsections that follow describe the scope of this document in specific areas.

1.2 *High-Level Purpose*—The purpose of this guide includes: (1) educating new users on laboratory informatics tools; (2) providing a standard terminology that can be used by different vendors and end users; (3) establishing minimum requirements for laboratory informatics; (4) providing guidance for the specification, evaluation, cost justification, implementation, project management, training, and documentation of the systems; and (5) providing a functional requirements checklist for laboratory informatics systems that can be adopted within the laboratory and integrated with existing systems.

1.3 *Laboratory Informatics Definition*—Laboratory informatics is the specialized application of information technology aimed at optimizing laboratory operations. It is a collection of informatics tools utilized within laboratory environments to collect, store, process, analyze, report, and archive data and information from the laboratory and its supporting processes. Laboratory informatics includes the effective use of critical data management systems, the electronic delivery of results to customers, and the use and integration of supporting systems (for example, training and policy management). Examples of

primary laboratory informatics tools include laboratory information management systems (LIMS), laboratory execution systems (LES), laboratory information systems (LIS), electronic laboratory notebooks (ELN), scientific data management systems (SDMS), and chromatography data systems (CDS).

1.4 *Scope Considerations when Selecting and Implementing Laboratory Informatics Solutions*—Many laboratories have determined that they need to deploy multiple laboratory informatics systems to automate their laboratory processes and manage their data. Selection of an informatics solution requires a detailed analysis of the laboratory's requirements and should not be a simple product category decision. Information technology (IT) representatives and subject matter experts (SMEs) who understand the needs of the laboratory need to be involved in the selection and implementation of a laboratory informatics system to ensure that the needs of the laboratory are met and IT can support it. Customers (internal and external) of laboratory information should also be included in the laboratory informatics solution design to ensure full electronic integration between systems.

1.5 The scope of this guide covers a wide range of laboratory types, industries, and sizes. Examples of laboratory types and industries include:

1.5.1 *General Laboratories:*

1.5.1.1 Standards (ASTM, IEEE, ISO) and

1.5.1.2 Government (EPA, FDA, JPL, NASA, NRC, USDA, USGS, FERC).

1.5.2 *Environmental:*

1.5.2.1 Environmental monitoring.

1.5.3 *Life Science Laboratories:*

1.5.3.1 Biotechnology and

1.5.3.2 Diagnostic.

1.5.4 *Healthcare and Medical:*

1.5.4.1 Bionomics/genomics,

1.5.4.2 Medical devices,

1.5.4.3 Pharmaceutical,

1.5.4.4 Veterinary,

1.5.4.5 Public health, and

1.5.4.6 Hospital.

1.5.5 *Heavy Industry Laboratories:*

1.5.5.1 Energy and resources,

1.5.5.2 Manufacturing and construction,

1.5.5.3 Materials and chemicals, and

¹ This guide is under the jurisdiction of ASTM Committee E13 on Molecular Spectroscopy and Separation Science and is the direct responsibility of Subcommittee E13.15 on Analytical Data.

Current edition approved Aug. 1, 2018. Published September 2018. Originally approved in 1993. Last previous edition approved in 2013 as E1578-13. DOI: 10.1520/E1578-18.

1.5.5.4 Transportation and shipping.

1.5.6 *Food and Beverage Laboratories:*

1.5.6.1 Agriculture,

1.5.6.2 Beverages,

1.5.6.3 Food, and

1.5.6.4 Food service and hospitality.

1.5.7 *Public Sector Laboratories:*

1.5.7.1 Law enforcement/forensic,

1.5.7.2 State and local government,

1.5.7.3 Education and nonprofits, and

1.5.7.4 Public utilities (water, electric, waste treatment).

1.6 *Integration*—The scope of integration covered in this guide includes communication and meaningful data exchange between different laboratory informatics tools and other external systems (document management, chromatography data systems, laboratory instruments, spectroscopy data systems, enterprise resource planning (ERP), manufacturing execution systems (MES), investigations/deviations and CAPA management systems), and other integrated business systems (for example, clinical or hospital environments) provide significant business benefits to any laboratory and is discussed at a high level in this guide.

1.7 *Life-Cycle Phases*—The scope of this guide is intended to provide an understanding of laboratory informatics tools' life cycle from project initiation point to retirement and decommissioning. This guide was designed to help newer audiences in understanding the complexity in the relationships between different laboratory informatics tools and how to plan and manage the implementation project, while seasoned users may use the different life cycles to maintain existing laboratory informatics tools. Integrating additional informatics tools to existing ones in today's evolving laboratory environment adds constraints that need to be considered. The life-cycle discussion includes both the laboratory informatics solution life cycle as well as the project life cycle.

1.7.1 The product life cycle encompasses a specific laboratory informatics system and the expected useful life of that system before it needs to be replaced or upgraded.

1.7.2 The project life cycle encompasses the activities to acquire, implement, operate, and eventually retire a specific laboratory informatics system.

1.8 *Audience*—This guide has been created with the needs of the following stakeholders in mind: (1) end users of laboratory informatics tools, (2) implementers of laboratory informatics tools, (3) quality personnel, (4) information technology personnel, (5) laboratory informatics tools vendors, (6) instrument vendors, (7) individuals who approve laboratory informatics tools funding, (8) laboratory informatics applications support specialists, and (9) software test/validation specialists. Information contained in this guide will benefit a broad audience of people who work in or interact with a laboratory. New users can use this guide to understand the purpose and functions of the wide variety of laboratory informatics tools as well as the interactions between these tools with external systems. The guide can also help prospective users in understanding terminology, configurations, features, design, benefits, and costs of these different laboratory informatics tools.

Individuals who are purchasing specific tools may also use this guide to identify functions that are recommended for specific laboratory environments. Research and development staff of different commercial laboratory informatics systems vendors may use the guide as a tool to evaluate, identify, and potentially improve the capabilities of their products. The vendors' sales staff may use the guide to represent functions of their laboratory informatics products to prospective customers in more generic and product-neutral terms.

1.9 *Out of Scope*—This guide does not attempt to define the boundaries of laboratory informatics, as they continue to evolve and blur between the different types of tools; rather, it focuses on the functionality that is provided by laboratory informatics as a whole.

1.10 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*²

E1394 Specification for Transferring Information Between Clinical Instruments and Computer Systems (Withdrawn 2002)³

E1947 Specification for Analytical Data Interchange Protocol for Chromatographic Data

E1948 Guide for Analytical Data Interchange Protocol for Chromatographic Data

E2077 Specification for Analytical Data Interchange Protocol for Mass Spectrometric Data

E2078 Guide for Analytical Data Interchange Protocol for Mass Spectrometric Data

E2369 Specification for Continuity of Care Record (CCR)

2.2 *CDISC Standard:*⁴

SEND Standard for Exchange of Nonclinical Data;

2.3 *CIDX Standard:*⁵

CIDX Chemistry Industry Data eXchange

2.4 *EPA Standard:*⁶

ESAR Environmental Sampling, Analysis and Results Data Standard

2.5 *FDA Regulation:*⁷

21 CFR Part 11 Electronic Records; Electronic Signatures, 62 FR 13464

² For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

³ The last approved version of this historical standard is referenced on www.astm.org.

⁴ Available from Clinical Data Interchange Standards Consortium (CDISC), 401 West 15th Street, Suite 800, Austin, TX 78701, <https://www.cdisc.org/>.

⁵ The CIDX Chem eStandards are available at OAGi, P.O. Box 4897, Marietta, GA 30061-4897, <http://www.oagi.org/>.

⁶ Available from United States Environmental Protection Agency (EPA), 1200 Pennsylvania Ave., NW, Washington, DC 20460, <http://www.epa.gov/>.

⁷ Available from U.S. Food and Drug Administration (FDA), 10903 New Hampshire Ave., Silver Spring, MD 20993, <http://www.fda.gov>.

FDA Data Integrity and Compliance with CGMP: Guidance for Industry

2.6 **HL7 Standards:**

Health Level Seven⁸ Standards

2.7 **ICH Standard:**⁹

ICH Quality Guideline Q9 Quality Risk Management

2.8 **IEEE Standards:**¹⁰

IEEE 1012 Standard for System, Software, and Hardware Verification and Validation

IEEE 1028 Standard for Software Reviews and Audits

2.9 **ISA Standard:**¹¹

ANSI/ISA-95.00.06 Enterprise-Control System Integration-- Part 6: Messaging Service Model

2.10 **ISO Standards:**¹²

ISO/IEC/IEEE 12207 Systems and software engineering -- Software life cycle processes

ISO/IEC 27000 Information technology -- Security techniques -- Information security management systems -- Overview and vocabulary

ISO/HL7 27932 Data Exchange Standards -- HL7 Clinical Document Architecture, Release 2

ISO/IEEE 11073-10101 Health informatics -- Point-of-care medical device communication -- Part 10101: Nomenclature

ISO/IEC/IEEE 26511 Systems and software engineering -- Requirements for managers of user documentation

ISO/IEC/IEEE 26512 Systems and software engineering -- Requirements for acquirers and suppliers of information

ISO/IEC/IEEE 29119-4 Systems and software engineering -- Software testing -- Part 4: Test techniques

ISO/IEC/IEEE 29119-5 Systems and software engineering -- Software testing -- Part 5: Keyword-Driven Testing

ISO/IEC/IEEE 29148 Systems and software engineering -- Life cycle processes -- Requirements engineering

2.11 **ISPE GAMP**¹³ Guides:

ISPE GAMP 5: A Risk-based Approach to Compliant GxP Computerized Systems

ISPE GAMP Guide: Records & Data Integrity

2.12 **MHRA Standard:**¹⁴

MHRA GxP Data Integrity Definitions and Guidance for Industry

2.13 **NCPDP Standard:**¹⁵

Batch Transaction Format

2.14 **NIST Standard:**¹⁶

NIST Cybersecurity Framework

2.15 **NRC Standards:**¹⁷

10 CFR Appendix B to Part 50 Quality Assurance Criteria for Nuclear Power Plants and Fuel Reprocessing Plants, 72 FR 49505

10 CFR Appendix E to Part 50 Emergency Planning and Preparedness for Production and Utilization Facilities, 80 FR 74980

10 CFR Appendix K to Part 50 ECCS Evaluation Models, 65 FR 34921

2.16 **PIC/S Standard:**¹⁸

PIC/S Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments

2.17 **Regenstrief Institute Standard:**

LOINC¹⁹ Logical Observation Identifiers Names and Codes

2.18 **SNOMED International Standard:**

SNOMED-CT Systematized Nomenclature of Medicine-Clinical Terms²⁰

2.19 **WHO Standard:**²¹

WHO Technical Report Series, No. 996, Annex 5 Guidance on good data and record management practices

3. Terminology

3.1 *Definitions*—This guide defines the majority of terminology used in the field of laboratory informatics. Users of this guide should request a terminology list from each vendor with a cross reference to the terms used in this guide.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *artificial intelligence, AI, n*—behavior by machines or computers versus the natural intelligence of humans and animals.

3.2.1.1 *Discussion*—In the computer science arena, any device that perceives its environment and takes action to maximize success in achieving a goal is exhibiting AI. Machine learning is an application of artificial intelligence that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

3.2.2 *chromatography data system, CDS, n*—computer system used to acquire, analyze, store, and report information from chromatographic instruments.

¹⁵ Available from National Council for Prescription Drug Programs (NCPDP), 9240 East Raintree Drive, Scottsdale, AZ 85260-7518, <https://www.ncdp.org/>.

¹⁶ Available from National Institute of Standards and Technology (NIST), 100 Bureau Dr., Stop 1070, Gaithersburg, MD 20899-1070, <http://www.nist.gov>.

¹⁷ Available from U. S. Nuclear Regulatory Commission (NRC), 11555 Rockville Pk., Rockville, MD 20852, <http://www.nrc.gov>.

¹⁸ Available from PIC/S Secretariat, 14 rue du Roveray CH, 1207 Geneva, Switzerland, <https://picscheme.org/>.

¹⁹ Registered trademark of and available from The Regenstrief Institute, Inc, 410 West 10th Street, Suite 2000, Indianapolis, IN 46202-3012, <http://loinc.org>.

²⁰ Available from SNOMED International, One Kingdom St., Paddington Central, London W2 6BD, United Kingdom, <https://www.snomed.org/>.

²¹ Available from World Health Organization (WHO), Avenue Appia 20, 1202 Geneva, Switzerland, <http://www.who.int>.

⁸ Registered trademark of and available from Health Level Seven (HL7) International, 3300 Washtenaw Avenue, Suite 227, Ann Arbor, MI 48104, <http://www.hl7.org/>.

⁹ Available from International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), ICH Secretariat, 9, chemin des Mines, P.O. Box 195, 1211 Geneva 20, Switzerland, <http://www.ich.org>.

¹⁰ Available from Institute of Electrical and Electronics Engineers, Inc. (IEEE), 445 Hoes Ln., Piscataway, NJ 08854-4141, <http://www.ieee.org>.

¹¹ Available from The International Society of Automation (ISA), 67 T. W. Alexander Drive, P. O. Box 12277, Research Triangle Park, NC 27709, <http://www.isa.org>.

¹² Available from International Organization for Standardization (ISO), ISO Central Secretariat, BIBC II, Chemin de Blandonnet 8, CP 401, 1214 Vernier, Geneva, Switzerland, <http://www.iso.org>.

¹³ Registered trademark of and available from International Society for Pharmaceutical Engineering (ISPE), 600 N. Westshore Blvd., Suite 900, Tampa, FL 33609, <http://www.ispe.org>.

¹⁴ Available from Medicines and Healthcare products Regulatory Agency, Her Majesty's Government, 10 South Colonnade, London E14 4PU, United Kingdom, <https://www.gov.uk/>.

3.2.3 *cloud computing, v*—term generally used to refer to software applications that are delivered as a software service through remote hosting using the public internet (public cloud) or within the users' network environment (private cloud).

3.2.3.1 *Discussion*—Essentially, the difference between cloud computing and traditional application deployment is that the application's users may not be responsible for the installation and maintenance of the computing infrastructure and application software.

3.2.4 *corrective and preventative action, CAPA, n*—CAPA applications are used to collect information, analyze information, identify and investigate product and quality problems, and take appropriate and effective corrective or preventive (or both) action to prevent their recurrence.

3.2.4.1 *Discussion*—Verifying or validating corrective and preventive actions, communicating corrective and preventive action activities to responsible people, providing relevant information for management review, and documenting these activities are essential in dealing effectively with product and quality problems, preventing their recurrence, and preventing or minimizing device failures.²²

3.2.5 *cybersecurity, n*—set of technologies, practices, and processes used to protect computers, networks, programs, and data from attack, damage, exploitation, and unauthorized access.

3.2.6 *data exchange standardization, n*—as defined by the International Organization for Standardization (ISO) in ISO/HL7 27932, the process of agreeing on standards, which represent the common language that allows the exchange of data between disparate data systems.

3.2.6.1 *Discussion*—The goals of standardization are to achieve comparability, compatibility, and interoperability between independent systems, to ensure compatibility of data for comparative statistical purposes, and to reduce duplication of effort and redundancies. A data standard often includes data elements, data element definitions, and such agreements as formats, message structures, and vocabulary. In the context of this paper, a standard is a specification or requirement and is not synonymous with a policy, procedure, guideline, framework, technique, or best practice. Adopting standards has the potential to improve interoperability and reduce costs by facilitating the ability of networked laboratories to coordinate activities during public health incidents where surge capacity may be required (for example, national response and readiness). Adopting standards may reduce the costs of laboratory informatics system implementation and vendor/developer support.

3.2.7 *data integrity, n*—extent to which data are attributable, complete, consistent, accurate, and reliable throughout the data life cycle.

3.2.8 *electronic document management system, EDMS, n*—computer system used to store, catalog, review/approve, retrieve, view, and print digital documents.

3.2.8.1 *Discussion*—Modern EDMS applications typically provide the ability to manage a document throughout its life

cycle, including document initiation, multiple review levels, version control, security, and archiving of historical versions of documents.

3.2.9 *electronic laboratory notebook, ELN, n*—software program designed to replace paper laboratory notebooks; an electronic system on which to create, store, retrieve, and share fully electronic records in ways that meet all legal, regulatory, technical, and scientific requirements.

3.2.9.1 *Discussion*—Laboratory notebooks, in general, are used by scientists, engineers, and technicians to document research, experiments, and procedures performed in a laboratory. A laboratory notebook is often maintained to be a legal document and may be used in a court of law as evidence. Similar to an inventor's notebook, the laboratory notebook is also often referred to in patent prosecution and intellectual property litigation.

3.2.10 *electronic signature, n*—electronic representation of a handwritten signature.

3.2.11 *enterprise resource planning, ERP, n*—computer system to integrate different types of data such as inventory levels, product orders, manufacturing capacity, inspection results, accounting data, and customer relationship management information from organizations within an enterprise (company), facilitating the flow of information between various business functions across a company as well as with outside business partners.

3.2.12 *good automated manufacturing practice forum, GAMP Forum, n*—volunteer group under the auspices of the International Society of Pharmaceutical Engineering (ISPE) for writing guidance for the validation of computerized systems used in the regulated portions of the pharmaceutical and allied industries and it is specifically designed to aid suppliers and users in the pharmaceutical industry.

3.2.13 *integration broker, n*—messaging application that can receive or extract data from a source system at the appropriate time, transform the data, and route the reformatted data to the target node.

3.2.13.1 *Discussion*—An integration broker application can also provide a repository for archiving, searching, and retrieving these messages.

3.2.14 *internet of things, IoT, n*—system of objects—computing devices, machines, objects, people, animals, and so forth—that can connect to a network and communicate among themselves, often without human intervention.

3.2.14.1 *Discussion*—An IoT device is an object operating within that system.

3.2.15 *laboratory execution system, LES, n*—computer system used in the laboratory at the analyst work level to aid in step enforcement for laboratory test method execution.

3.2.15.1 *Discussion*—LES focus on step execution of defined laboratory test methods. The LES is typically used in quality control laboratories that have defined test methods. The functionality of a LES and a laboratory information management system (LIMS) overlaps in the areas of result entry, instrument integration, and specification flagging. Deployment

²² For additional information, visit the FDA's CAPA page at <http://www.fda.gov/>.

options include LES and LIMS systems deployed as an integrated solution, LIMS-only, or LES-only (for limited functions).

3.2.16 *laboratory informatics, n*—term used to describe the specialized application of information technology aimed at optimizing laboratory operations.

3.2.16.1 *Discussion*—That technology includes informatics tools used within laboratory environments to collect, store, process, analyze, report, and archive data and information from the laboratory and supporting processes. Laboratory informatics includes the effective use of critical data management systems, the electronic delivery of results to customers, and the use and integration of supporting systems (for example, training and policy management). Examples of primary laboratory informatics tools include, LIMS, LES, CDS, ELN, laboratory information systems (LIS), and scientific data management systems (SDMS).

3.2.17 *laboratory informatics tools configuration, n*—refers to the process of changing the functions of any laboratory informatics tool to match the business processes used in a particular laboratory, and it does not involve the use of writing software code either via a recognized software language or a language provided by the informatics application supplier.

3.2.17.1 *Discussion*—This is a GAMP Category 4 software and is defined as “Configured software including, LIMS, SCADA, DCS, CDS, etc.” Such configuration typically involves using an interface provided by the vendor to enter information that describes the types of samples, analytical methods, specifications, and so forth, used in the laboratory. It may also involve the configuration of options and businesses rules within the tool.

3.2.18 *laboratory informatics tools customization, n*—refers to the process of changing the functions of any laboratory informatics tool to match the business processes used in a particular laboratory.

3.2.18.1 *Discussion*—This is different from the previously mentioned tools configuration in that customization involves writing software code either via a recognized software language or a language provided by the informatics application supplier. This is a GAMP 5 software category. Such customization typically involves adding tables, modifying table structures, and writing code or programs to alter the behavior of any laboratory informatics tool.

3.2.19 *laboratory information management system, LIMS, n*—(1) computer software and hardware that can acquire, analyze, report, and manage data and information in the laboratory; (2) computer software that is used in the laboratory for the management of samples, test results, laboratory users, instruments, standards, and other laboratory functions such as invoicing, plate management, product/material stability programs, and work flow automation; and (3) a class of application software which handles storing and managing of information generated by laboratory processes.

3.2.19.1 *Discussion*—These systems are used to manage laboratory processes, including master data definition, sample management and chain of custody, work assignment, instrument and equipment management, standard and reagent

management, scheduled sample collection and testing, result entry, capture of results from instruments, result review, reporting, trending, and business rule enforcement. These systems interface with laboratory instruments (for example, CDS and other information systems such as ERP, LIS, or manufacturing execution systems [MES]). A LIMS is a highly flexible application, which can be configured or customized to facilitate a wide variety of laboratory workflow models.

3.2.20 *laboratory information system, LIS, n*—class of application software that supports clinical laboratories by helping laboratory personnel manage the quality and integrity of test samples, departmental workflow functions, result review processes, reports of finalized results, interpretations, and diagnoses.

3.2.20.1 *Discussion*—These systems often interface with instruments and other information systems such as hospital information systems (HIS). A LIS is a highly configurable application and often includes laboratory-specific electronic medical records, direct clinician access via secure web connections, billing modules for laboratories performing commercial testing, sophisticated interface engines for routing orders and results to external systems, and on-board image archival systems for pathology images. Patient confidentiality and HIPAA requirements define unique security functionality for a LIS. The College of American Pathologists (CAP) publishes LIS product guides²³ that list current LIS in the market.

3.2.21 *lean laboratory, n*—set of management and organizational processes that enables efficient testing flow, leveled workloads, visual work assignment and tracking, and the elimination of waste.

3.2.21.1 *Discussion*—Lean laboratory designs yield productive, high-quality laboratory environments that are sometimes supported by laboratory informatics tools.

3.2.22 *mapping tools, n*—graphical data mapping, conversion, and integration applications that map data between any combination of XML, database, flat file, EDI, Excel (OOXML), XBRL, or web service, or both, then transforms data or autogenerates data integration code for the execution of recurrent conversions.

3.2.23 *master data, n*—represents the business objects which are agreed on and shared across the enterprise.

3.2.23.1 *Discussion*—It can include relatively static reference, transactional, unstructured, analytical, and hierarchical data, as well as associated metadata. Examples of master data include product specifications, test method steps (to capture intermediate and final results), laboratory calculations, instrument information, and standard and reagent information.

3.2.24 *metadata, n*—(1) data about data and (2) information that describes another set of data.

3.2.24.1 *Discussion*—Additional information about the data that provides context and meaning, including how, when, and by whom it was collected, and its relationship to the subject or test. Metadata in any laboratory informatics tool’s context

²³ For additional information, visit CAP’s product guide page at <http://www.captodayonline.com>.

typically includes all data that supports a test result that is recorded in this tool. For example, a pH test includes a pH result that can be supported by metadata, including what instrument was used, what the calibration date of the instrument was, what standard buffer solutions (reagents) were used to calibrate the pH probe sensor, the expiration dates for the standard solutions, and the temperature of the solution at time of measurement.

3.2.25 *sample registration, n*—process of recording incoming sample information in a given laboratory informatics tool.

3.2.26 *scientific data management system, SDMS, n*—computer system used to capture, centrally store, catalog, and manage data generated in a laboratory environment.

3.2.26.1 *Discussion*—These data are then available for re-use and integration with other laboratory informatics systems. An example of an SDMS is an electronic repository for reports from laboratory informatics systems. The SDMS may include raw data file storage and archiving of data. It may also provide e-signature functionality for review/approval.

3.2.27 *spectroscopic data systems, n*—computer systems used to collect, process, visualize, interpret, store, and report information from spectroscopic instruments.

4. Significance and Use

4.1 *Relevance*—This guide is intended to educate the intended audience on many aspects of laboratory informatics. Specifically, the guide may:

- 4.1.1 Help educate new users of laboratory informatics;
- 4.1.2 Help educate general audiences in laboratories and other organizations that use laboratory informatics;
- 4.1.3 Help educate instrument manufacturers and producers of other commonly interfaced systems;
- 4.1.4 Provide standard terminology that can be used by laboratory informatics vendors and end users;
- 4.1.5 Establish a minimum set of requirements for primary laboratory informatics functions;
- 4.1.6 Provide guidance on the tasks performed and documentation created in the specification, evaluation, cost justification, implementation, project management, training, and documentation of laboratory informatics; and
- 4.1.7 Provide high-level guidance for the integration of laboratory informatics and other software tools.

4.2 *How to be Used*—This guide is intended to be used by all stakeholders involved in any aspect of laboratory informatics implementation, use, or maintenance.

4.2.1 It is intended to be used throughout the laboratory informatics life cycle by individuals or groups responsible for laboratory informatics implementation and use, including specification, build/configuration, validation, use, upgrades, and retirement/decommissioning.

4.2.2 This guide also provides an example of a laboratory informatics functional requirements checklist that can be used to guide the purchase, upgrade, or development of a laboratory informatics system.

5. Elements of Laboratory Informatics

5.1 *Laboratory Informatics Elements Overview*—Laboratory informatics is used to help laboratory personnel

better collect, process, analyze, report, store, and share the data and information derived from the laboratory and its supporting processes. These processes are often an integral part of a laboratory's workflow and include activities such as registration of samples or experiments, or both, assignment of tests, entry of results, review and approval of results, and reporting. Laboratory informatics' scope encompasses multiple technical solutions or systems that are responsible for streamlining these and other laboratory processes. Laboratory informatics is not solely about software managing laboratory data; it has many elements, some of which integrate or cross over with business management and other third-party tools. Those elements are also becoming increasingly complex, both in functionality and interoperability. Outside of standard laboratory information management systems (LIMS) and laboratory information systems (LIS), elements such as field data capture systems, advanced analytics tools, and artificial intelligence continue to shape the field of laboratory informatics. The division between these and other system categories continues to soften as functionality continues to be added to each of them. LIMS were originally created to address laboratories' need to manage laboratory operations and data, provide traceability for all laboratory samples and equipment, and ensure that laboratory procedures are followed. Electronic laboratory notebooks (ELNs), on the other hand, were originally created to meet scientists' need to document their experimental design, execution, and conclusion in an electronic format instead of in a paper notebook. The scientific data management system (SDMS) was created to provide a repository of all scientific data files and results regardless of instrument type. The current definitions of each of these system categories are far more encompassing and continue to evolve as the boundaries between categories continue to blur. That blurring of laboratory informatics elements, as well as their potential integration with enterprise elements—both within organizations and with customers of laboratory information—are illustrated in Fig. 1. Laboratory informatics and all it encompasses is shown with the large yellow circle on the left, while the internal business systems that support laboratories are found associated with the blue circle on the right. Surrounding both is a bubble representing third-party interactions with both laboratory generated data and business data. The figure highlights the wide variety of crossover and interactions that can occur both within and external to laboratory informatics. Laboratory informatics applications are also taking on some of the functionality of internal business systems. (See 7.6–7.11 for more information on integration cases.)

5.1.1 *Core Systems*—These laboratory systems most often provide the outward face of laboratory informatics and include LIMS, LIS, laboratory execution systems (LES), ELN, SDMS, and chromatography data systems (CDS). Not all systems will necessarily appear in a laboratory together; some are more typical to certain laboratory types than others. However, they usually play a key role in a laboratory's research or analysis activities, or both, and represent the key software systems with which laboratory personnel and customers of the laboratory may interact. From research samples and clinical specimens to

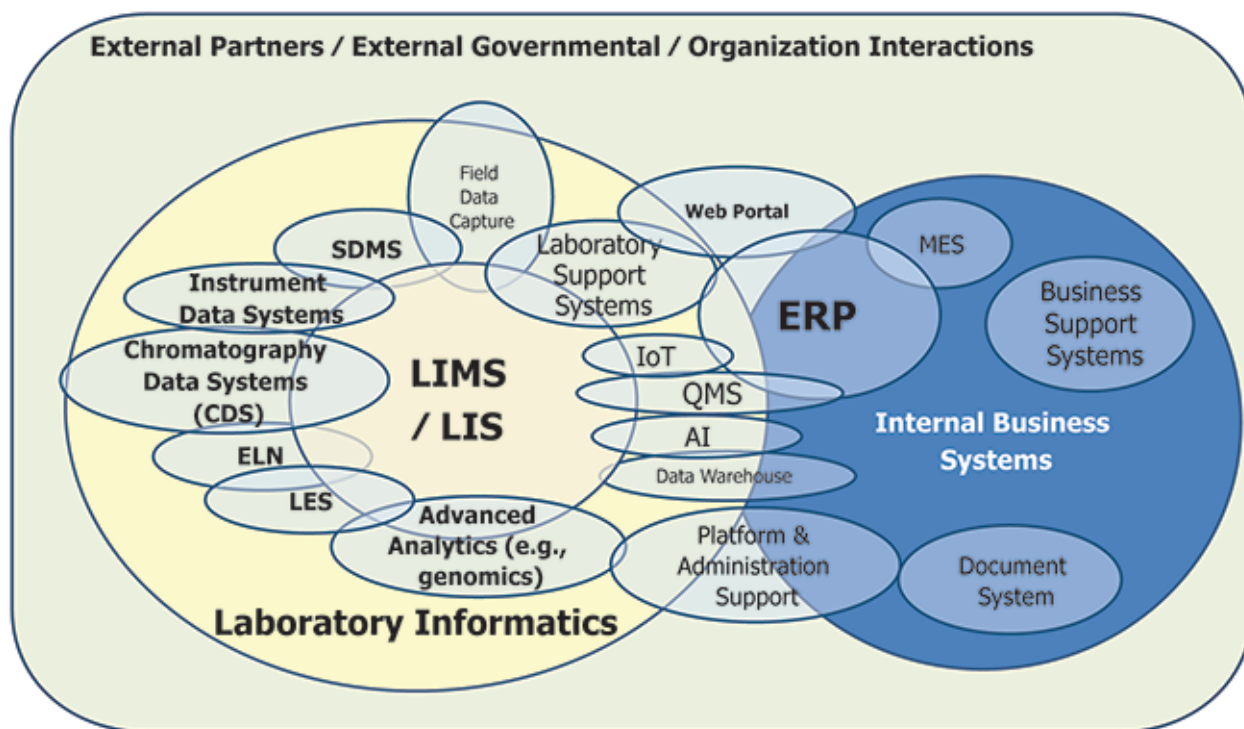


FIG. 1 Laboratory Informatics System Integration Concept Model

outlined experiments and raw instrument data, these core systems fill a vital role in the laboratory informatics sphere.

5.1.1.1 LIMS, LIS, and LES are alike in many regards in that they all act as core systems in a laboratory and handle data capture, analysis, review, storage, and reporting. These systems integrate in variable degrees with analytical instruments, automated tools, and other software systems, and they provide certain levels of regulated, industry-standard security for the data generated and transferred from its integrations. However, these systems also have fundamental differences that place them in specific use cases. A LIMS has been traditionally used to process and report on batches of samples from research, quality control, and manufacturing laboratories, all of which handle mostly anonymous, complex laboratory data. A LIS has normally been used in the clinical context of specimens and patients, and a LES is most often adopted in automated and regulated manufacturing environments where quality control, process control, test step execution, and instrument interface and calculation validation support laboratory testing.

5.1.1.2 An ELN largely serves as an electronic replacement for the traditional paper laboratory notebook associated with scientists and technicians in research-driven environments. The ELN may be tailored to the individual researcher, large collaborative research efforts, or both. It may also be designed to manage the activities related to a specific scientific discipline or application, or it may be cross-disciplinary, supporting data of all types. Traditionally used to document experiments and analysis, and act as intellectual property protection, the paper laboratory notebook has fallen out of favor with some laboratories that prefer the ELN's ability to improve search, support collaboration, and limit siloed data. Additional features of an

ELN include data import, content linking, preformatted and customizable templates, and messaging.

5.1.1.3 An SDMS is designed primarily to consolidate data and manage knowledge-based assets. The SDMS has typically excelled at handling unstructured files such as images, spreadsheets, raw instrument files, and PDFs. A set of agreed-upon rules in the system dictate how incoming data is processed and structured in the SDMS, acting as a gatekeeper for what is captured and how (including the application of metadata). An SDMS can be integrated with a LIMS, ELN, and so forth, to create a common repository for a laboratory's data, which can then be further associated with specific projects, experiments, or locations, or combinations thereof, or its data can simply be archived for long-term storage.

5.1.1.4 A CDS is designed for collecting, processing, and analyzing samples run on instruments managing chromatography techniques such as high-performance liquid chromatography (HPLC), ion chromatography (IC), gas chromatography (GC), size-exclusion chromatography (SEC), and affinity chromatography. The CDS typically consists of a combination of hardware and software connecting the instrument to the system and is computationally intensive, rapidly generating large data sets in laboratory environments. Complex algorithms and mathematical transformations of data can be performed within a CDS, which directly supports a wide range of chromatography instruments with bidirectional control of instrument settings (that is, temperature, pressure, and detector wavelength). The CDS typically provides sample handling (auto samplers), auto injection, mobile phase controls (temperature/pressure), detector control (wavelength), data collection (data points from one or more detectors), data analysis (for example, calibration

curves, peak detection, and integration), reporting, and audit trail support. The CDS can be deployed as a standalone system or in larger configurations supporting multiple instruments, sites, and geographic regions. The CDS is typically interfaced to other laboratory informatics tools (that is, LIMS, ELN, and SDMS) in which sample information is passed between the LIMS and the CDS, and the CDS test results (that is, peak areas) are passed back to the LIMS for final reporting. The CDS can also integrate with multiple analytical techniques in which data are passed between different instruments [such as a liquid chromatography-mass spectrometry (LC-MS) instrument]. The software is typically treated as a separate laboratory informatics element with its own IT infrastructure, including data acquisition modules (to attach to chromatography instruments) and administration, configuration, and security access controls.

5.1.1.5 **Bioinformatics and genomic applications** are used to collect and analyze biological data, including genomic data. Genomic instruments generate much larger amounts of data compared to traditional laboratory instruments. Genomic data is often analyzed using specialized, proprietary bioinformatics algorithms.

5.1.2 **Instrument Data Systems**—Some laboratory systems are purpose-made for specific instruments, with the CDS being a common case. Other examples include data acquisition systems (DAQ) for calibration equipment and digital oscilloscope software. These systems excel at “talking” with a specific instrument, capturing the data, and providing custom analysis tools related to the captured data. In the case of chromatography, a CDS will often both control a chromatography or spectroscopy instrument and provide a visual report of the chromatogram with contextual meaning. DAQ for calibration instruments may include additional functionality such as real-time data visualization and certificate generation.

5.1.3 **Advanced Analytics Tools**—This element represents a broad category of advanced analytical tools used in laboratories. Examples include the scientific field of genomics in which advanced tools are used to study complete genomes (genetic material within an organism). High-throughput genome analyzer instruments automate process steps to provide high-volume parallel operations that combine DNA molecules, primers, polymerase amplification, imaging, and computational functions to yield low-cost DNA sequencing outputs. Advanced tools are also used to automate DNA sequence assembly (reconstruction of an original DNA sequence). Advanced annotation tools/instruments are used to annotate DNA sequences by identifying portions of the genome that do and do not code for proteins, including supporting biological information.

5.1.4 **Field Data Capture Systems**—Field data capture and analysis is important to several industries, including many which do their own laboratory testing. Along with its base functionality, these systems are also often capable of integrating supervisory control and data acquisition (SCADA)-related infrastructure data (for example, from a wastewater or oil and gas stream) with manually collected field worker data to improve test data analysis and decision support. They also

typically allow native mobile apps to integrate, improving data analysis and reducing manual data input time.

5.1.5 **Laboratory Support Systems**—These systems are typically ancillary to the core systems, filling in functionality gaps that the core systems do not provide. Quality management software is one such example, used to formulate quality policy and objectives, standard operating procedures, and the required records used for the quality certification process. Middleware that can handle auto-validation of samples represents another example. The following may all be considered part of this category:

5.1.5.1 **Artificial intelligence (AI) tools and algorithms** are being used by researchers to inspect data better and make discoveries in the laboratory, while laboratory tools such as freezers, incubators, and air-cleaning systems are becoming “smarter” with added sensors that can feed data to one or more software systems for monitoring.

5.1.5.2 **Batch and lot management tools** assist with the creation, genealogy, review, and disposition of samples en masse, in the process assigning the same properties and process tracking for increased efficiency.

5.1.5.3 **Capacity planning and laboratory scheduling** improves the efficiency of laboratory workflow, allowing laboratory personnel to use better the time and resources of available personnel and equipment. Such tools take into account scheduled instrument maintenance and planned time off of researchers.

5.1.5.4 **Compliance management**, whether embedded in a core system or installed as a support system, helps keep laboratories of all types on track with complying with government regulations. Aside from the typical audit trails and electronic signatures, compliance management tools also may assist with risk assessments, business governance, and contractual obligation management.

5.1.5.5 **Data integrity** is a core competency and expectation of laboratory informatics solutions. Data integrity includes the maintenance and assurance of the accuracy and consistency of data over its entire life cycle. Data integrity is a critical design requirement that touches every element of laboratory informatics that stores, processes, or retrieves laboratory data. (See Section 10 for more on data integrity.)

5.1.5.6 **Human resource management** encompasses the activities associated with managing the data surrounding laboratory personnel. This includes the management of analyst training records, qualifications, certifications, and performance.

5.1.5.7 **Instrument and equipment management tools** help to determine on- or offline status, assist with calibration management, update service and preventative maintenance schedules, and present the qualification status of instruments for proper safety and functionality.

5.1.5.8 **Instrument data capture and control** can provide some of the greatest benefits to both laboratory efficiency and data quality and integrity. Traditionally, laboratory informatics tools have allowed direct data capture from simple instruments, indirect data capture through instrument data file parsing, and, less frequently, bi-directional control of simple and file-based instruments. Although it is common for laboratory informatics

solutions to have integrated instrument integration tools, there are also standalone instrument integration systems that may act as a data capture hub for the laboratory by connecting many instruments to a single application. These hub systems may then be interfaced to laboratory informatics solutions to provide data from multiple instruments via a single system-to-system interface. Instrument integration may also be available as a service, hosted in the cloud or on-premises, and newer instrumentation and devices are available that take advantage of the internet of things (IoT) web connection capabilities that share their data.

5.1.5.9 Inventory management involves the handling of controlled substances, reagents, and retention samples, and it may also handle stability analysis.

5.1.5.10 Investigation management is important to researchers and manufacturers in the pharmaceutical, medical device, and biologics industries as part of a regulated process that strives to ensure the safety and compliance of the components, containers, in-process materials, and finished products of those industries. This includes the management of the FDA's out-of-specification (OOS) requirements,²⁴ internal out-of-trend (OOT) requirements, process validation efforts, and any corrective and preventative action (CAPA) efforts related to investigations.

5.1.5.11 Process improvement is generally thought of as a series of discussions and changes towards making the laboratory leaner, often through the concept of continuous improvement. However, visual project management “dashboards” that collate performance data and visually present it may be integrated with other laboratory informatics tools.

5.1.5.12 Scheduled event management is enhanced through the use of event management and planning software or functionality. In particular, laboratories with shifts or rotations that move laboratory personnel around in various sections benefit from tools that assist with routine and nonroutine work.

5.1.5.13 Scientific data management does not necessarily need to be limited to an SDMS. Other core, instrument data, and standalone laboratory systems are designed to manage a wide variety of instrument raw data files, documents, and experiments for the purpose of improved data homogenization, sharing, and mining.

5.1.5.14 Standard and reagent management encompasses verifying the storage, quantity, on-order, and expiry status of a wide variety of inventoried additives, chemicals, and calibration standards.

5.1.5.15 Statistical trending and control charting is part of the data analysis and visualization efforts of a software system. These tools allow laboratory personnel to improve decision making and process development, among other things.

5.1.5.16 Systems integration functionality allows laboratory informatics systems, business management systems [for example, enterprise resource planning (ERP)], manufacturing systems [for example, manufacturing execution system (MES)], process management tools (for example, SCADA),

document management systems, statistical analysis tools, and more to connect with each other and limit the amount of “siloes” or isolated data in an organization.

5.1.6 *Artificial Intelligence*—The application of AI in clinical, research, and industrial laboratories is still in relative infancy, though laboratory personnel and researchers alike are increasingly finding ways to integrate AI into laboratory workflow. With AI's ability to “learn” from user interactions, data feeds, and more to improve or make insights into laboratory processes and data, the technology will likely continue to find ways into the laboratory. Machine learning tools that analyze structured data and natural language processing tools that extract information from unstructured data are already finding their way into some clinical laboratory workflows.

5.1.7 *Platform and Administration Support*—Platform and administration support tools not only help laboratory informatics end users develop and maintain important documentation, they also allow systems administrators and IT personnel to provide a better, more secure experience. These tools may be built into a core system or act as standalone components of an interconnected web of systems. Functionality among these tools includes document management (for standard operating procedures, specifications, training material, and so forth), configuration management (for handling all aspects of master data, electronic signatures, and so forth), system validation and commission (for requirements, specifications, test scripts, regression and stress tests, and so forth), system administration, and electronic security and privacy management.

5.1.8 *Data Warehouse*—An SDMS is a type of data warehouse in which a series of agreed-upon rules dictate how the incoming data is processed and structured. But other forms of data warehousing exist as well, which both clinical and nonclinical laboratories alike are using it viably. The late-binding or “data lake” approach to data storage represents a more modern architecture that takes in data of various structural states, tags it with appropriate metadata, and flattens it, leaving it alone otherwise until queried or used, or both. These data storage systems may be hosted locally or offered as a service over the cloud.

5.1.9 *Web Portals*—These tools allow third parties to access, in a controlled fashion, one or more laboratory informatics data repositories or data management systems. They are often part of an existing system such as a LIMS but may also be standalone with a connection to one or more databases. The level of access granted to individuals or groups are controlled by the web portal's administrative tools.

5.2 *Laboratory Informatics Systems Evolution*—Laboratory informatics systems have evolved over time, with developers adding and expanding functions as capabilities and needs change. LIMS and instrument data systems such as CDS began by performing simple laboratory functions. Over time, additional software tools entered the laboratory, and existing software products added functionality. Fig. 2 shows a timeline for the development of software products designed to meet the needs of the laboratory community. The expanding breadth of tools available illustrates the increased functionality and complexity of laboratory informatics solutions. The laboratory

²⁴ See the Food and Drug Administration's “Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production,” available via the FDA at <https://www.fda.gov/>.

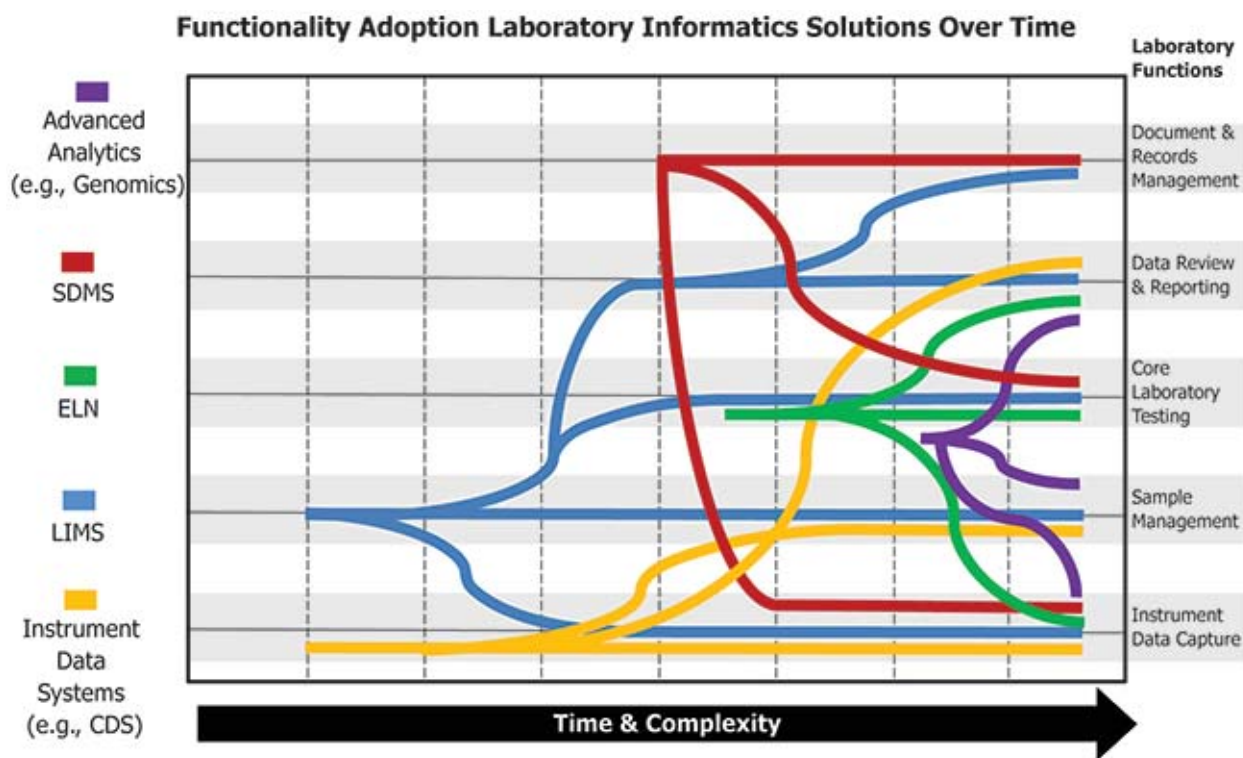


FIG. 2 Laboratory Informatics Systems Evolution

informatics solutions illustrated in this figure are examples, however, and do not imply that these are the only tools available.

5.3 *Laboratory Informatics Functions*—Laboratory informatics systems contain a rich set of functions. In Fig. 3, those functions using the primary functional categories that can be found in laboratory informatics systems are illustrated. At the center of these laboratory informatics functions is core laboratory testing (items labeled as C-x), including sample/experiment registration, sample management, testing, result review, sample approval, and reporting. Core laboratory testing is then supported by extended functions (items labeled as E-x), including planning and scheduling, instrument management, reagent and standard management, material disposition, experiment approvals, reporting, and trending. Finally, those core and supporting functions are propped up by laboratory informatics platform and administration support functions (items labeled as S-x and D-x), including master data management. Note that the labels assigned to these various sections directly tie into the items found in the requirements checklist of Appendix X1.

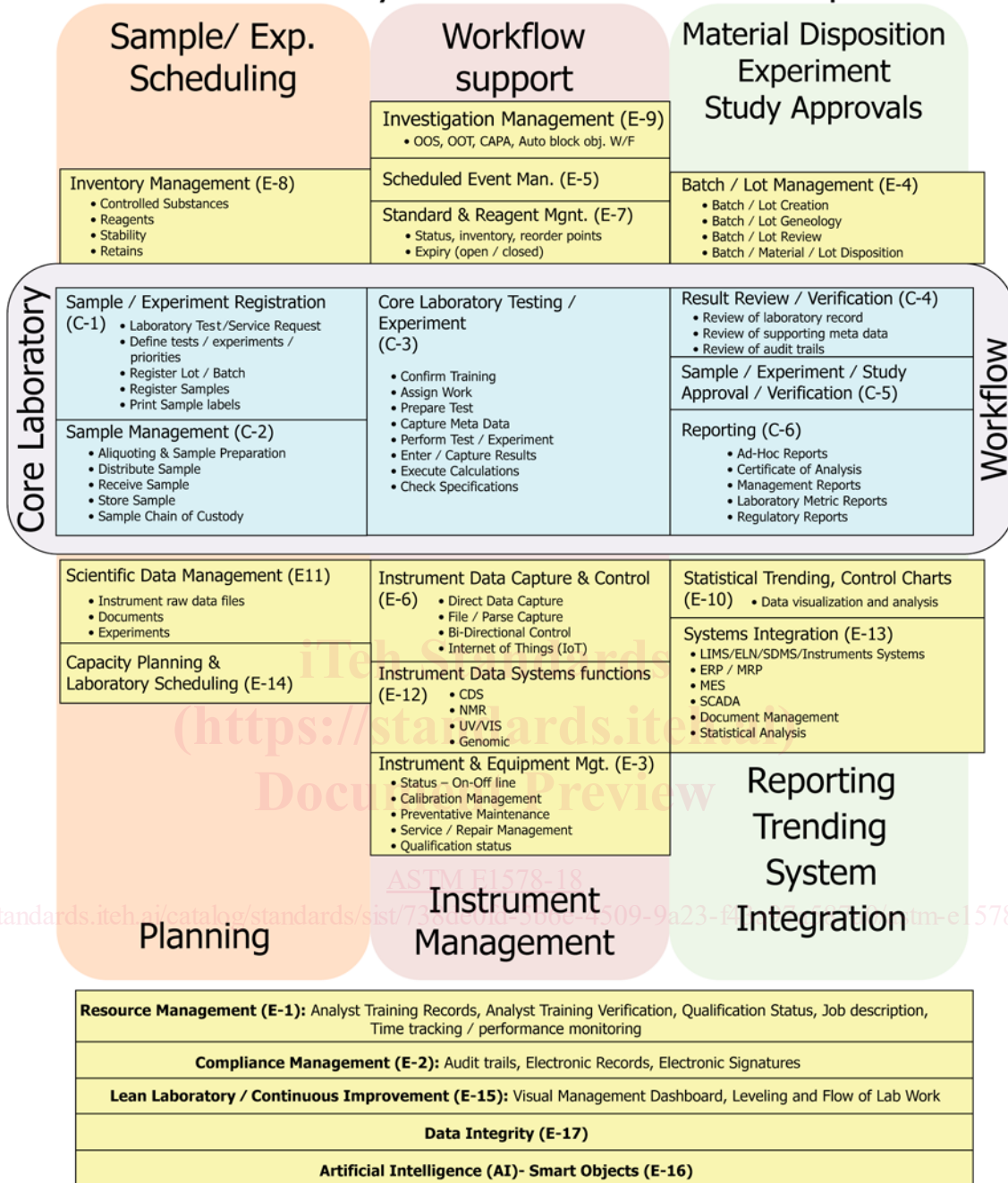
5.4 *Laboratory Informatics Functions for Specific Industries*—Laboratory informatics encompasses many technical solutions such as LIMS, LIS, ELN, and CDS, which aim to improve laboratory operations. However, not all laboratory workflows are the same. Laboratories in specific industries may require additional functionality to meet their workflow requirements. Some laboratory informatics systems vendors may design their solutions to meet the needs of many laboratory types, while others may design them for a particular industry segment with particular functionality needs. An environmental laboratory, for example, may require tracking of

sample containers, processing of samples in batches with control samples, instrument integration, multiple levels of review, and specific reporting requirements. In Fig. 4, some of the additional functions that may be required to address the needs of specific laboratory types are illustrated. The functions illustrated are over and above the basic laboratory workflow and are by no means an exhaustive list; these are merely examples of possible additional functionalities.

5.5 *Laboratory Informatics Implementation and Technological Considerations*—As with most any other software system, implementing laboratory informatics software in a laboratory requires at least a minimum set of plans and considerations. What operating systems are supported? What instruments shall be connected? Does the system need to be customized frequently? These types of questions quickly multiply as system complexity increases, operational needs expand, and integration needs become more complex. This line of questioning can become so complicated that functional requirements checklists are required to make sense of how a laboratory informatics solution fits into organizational workflow. As such, gaining a clearer understanding of a laboratory informatics solution’s life cycle, as well as the technological elements that go into its development and implementation, is vital to its successful operation in the laboratory.

5.5.1 The implementation life cycle of laboratory informatics systems is not unlike other software systems, though it comes with its own intricacies, including regulatory and “best practice” considerations specific to clinical, research, and other types of laboratories. The laboratory informatics system life cycle is described in more detail in Section 8.

Laboratory Informatics Function Map



Platform & Administration Support

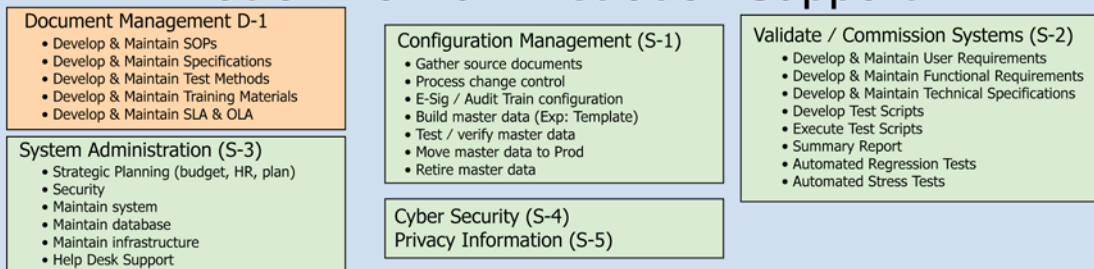


FIG. 3 Laboratory Informatics Functions

Additional Functions Required by Laboratory Types		Laboratory Functional Area					
		Sample / Experiment Management	Lab Testing	Review, Verification Approval	System Integration	Reporting and Trending	Other Functions
Laboratory Type	General Laboratories	Chain of custody	Batch Processing of Samples including QA		ELN, SDMS, instrument integration	Electronic Reporting of QC Results	
	Environmental	Container management, Chain of custody, regulatory compliance - significant auxiliary data	Batch Processing of Samples including QA	Multiple levels of review and approval	GIS integration	Electronic Reporting of QC Results by batch	Method QA/QC, Calculation validation, electronic data delivery
	Public Health Sector- (clinical microbiology and chemical)	Chain of custody, electronic test ordering, Emergency response, public health surveillance	Batch Processing of Samples including QA, agent regulations	Multiple levels of review and approval, CLIA certification	Emphasis on surveillance at multiple state and federal levels	Electronic Reporting multiple data formats, vocabulary, and secure transport	Clinical and non-clinical, sample centric and patient centric
	Life Sciences	Controlled substances Stability samples	Uniformity Calculations Product Specifications	Multiple levels of review and approval	ERP integration	SPC and Process variability analysis	Instrument maintenance and calibration tracking
	Food and Beverage	Lot geneology	Product Specifications		MES and ERP integration	SPC and Process variability analysis	Exception Reporting
	Heavy Industry	Automatic scheduling, Lot Management	Product Specifications		Process Control / PM integration	SPC and Process variability analysis	
	R&D	Experiment sharing	Method versioning		ELN, SDMS, instrument integration	Experiment Conclusions, Technical Reports	Expanded Search capability
	Medical Laboratory	Embedded LIS Standalone LIS Niche LIS Patient centric			interface engines to external systems	Unique functions Diagnostic and Surveillance	electronic medical records; direct clinician access, billing modules

FIG. 4 Laboratory Informatics Functions for Specific Industries

5.5.2 The technological elements of laboratory informatics systems and their design and deployment are worth noting. Technologies such as cloud computing and application programming interfaces (API) are continuing to shape system development and implementation across a wide variety of laboratory types. Key factors such as hardware infrastructure, application platform, and integration needs further drive decisions of how a system should be developed and deployed, thus influencing the use of technological elements towards those goals. The following may all be considered technological

elements relevant to laboratory informatics systems' development and deployment:

5.5.2.1 APIs allow software components to communicate clearly with each other, improving the way laboratory informatics applications are built and used. They not only allow core developers to simplify programming but also allow end users who need to customize aspects of how the software is used (and outputs data) a simpler way to do that. Some software vendors may ship their laboratory informatics offering with an API or software development kit (SDK) so the end user

can simplify how the software exchanges data with or calls background services from other systems in the laboratory, as well as customize the software in other ways.

5.5.2.2 Artificial intelligence (AI) and its application in the laboratory continue to grow with the use of smart objects, machine learning applied to laboratory data, and monitoring of laboratory processes and resources. See Section 11 for additional details on this subject.

5.5.2.3 Authentication mechanisms and protocols such as single sign-on (SSO) and OAuth are important to the security of a laboratory informatics application, and the decision of which to use is not taken lightly. These mechanisms and protocols help define how information is securely stored, accessed, and propagated. In some cases, regulations such as the Health Insurance Portability and Accountability Act (HIPAA) may further dictate selection, particularly in cases in which electronic-protected health information (ePHI) is involved.

5.5.2.4 Business intelligence (BI) and big data technology components increasingly find their way into or tangentially connected to laboratory informatics. This includes data warehousing, data mining, data analytics, data structuring, and web portal components that aid in the processing, managing, visualization, and dissemination of all sorts of data.

5.5.2.5 Cloud computing and storage offers a virtualized alternative to how laboratory informatics applications and data are leveraged. Whether because of a lack of information technology resources and internal expertise or a desire to reduce operating costs, informatics vendors and businesses alike are turning software into a service, accessible via the internet or over a private network. This technology has benefits and drawbacks, though it has also become more refined over the years. (See 7.1 for more.)

5.5.2.6 Data migration tools represent another important component of laboratory informatics deployment; many systems are not put into operation with a blank database. These tools are often standalone, meaning they do not become a permanent part of the overall informatics infrastructure, and they also may have data cleaning and transformation functionality to ensure data is in the right form when it arrives to its destination.

5.5.2.7 IoT devices have made their way to laboratories, promising greater productivity and insights. Collections of networkable sensors and devices give laboratories the option to interconnect, monitor, and make insights into various data streams, whether located in-house or thousands of miles away. However, implementing IoT in the laboratory comes with its own considerations, including communication methods, security, and quality of data. (See 7.2 for more.)

5.5.2.8 Software or technology stacks are essentially a collection of software subsystems or components that allow for the development and operation of a complete laboratory informatics solution. These subsystems and components are often chosen together frequently because of a solid history of integrating well, as well as the availability of abundant documentation to support the implementation of the stack. A

bioinformatics application, for example, may be built on a LAMP stack that includes Linux, Apache HTTP Server, MySQL, and Python.

5.6 *Laboratory Informatics: Standardization, Authentication, and Integration*—Several trends in laboratory informatics development, implementation, and use have slowly emerged, with a reasonable chance of impacting how laboratories manage data and its security. The concepts of data standardization, blockchain-based authentication, and end-to-end data integration are briefly described here.

5.6.1 Laboratory Informatics data standardization efforts have been active across multiple years and in different analytical techniques and industry segments. The adoption of data standards covering different laboratory instruments, different computer architectures, databases, and operating systems across different industry segments and laboratory types has been a slow tedious process. The demand and requirements for improved data-sharing mechanisms have expanded, driving further discussion and work efforts concerning electronic data formats. Data sources across a single scientific enterprise may number in the tens, making data sharing more difficult across hundreds of other such enterprises, many of which may include older or proprietary data sources that are not easily used by other entities. The durability of analytical data over long time periods is a growing concern. Data sources may be unorthodox too, such as from social media, focus groups, or regulatory filings in addition to traditional laboratory instrument data outputs. Emphasis on better managing, sharing, and mining this data has grown considerably, such that funding agencies, data organizations, standard organizations (like ASTM) and academic publishers are beginning to provide recommendations and requirements for shared laboratory data and how it is formatted, stored, and publicly accessed. Standardized electronic formats are important to laboratories of all types, especially public health and pharmaceutical laboratories. Groups such as Observational Health Data Sciences and Informatics are promoting data standardization efforts and data models such as the Observational Medical Outcomes Partnership (OMOP) Common Data Model²⁵ in recognition of this growing need. The OMOP data model attempts to transform, characterize, and analyze a wide variety of varying data sources associated with health care delivery and research “to ensure that research methods can be systematically applied to produce meaningfully comparable and reproducible results.”²⁶ Ongoing ASTM efforts supporting laboratory data standards include the Analytical Information Markup Language (AnIML),²⁷ an open XML-based format for analytical and biological data. The format supports data from many analytical and laboratory process techniques, enabling the creation of generic tools that can handle data from any instrument. Industry-specific consortiums (and their associated test method

²⁵ Available from Observational Health Data Sciences and Informatics (OHDSI), 622 W. 168th St., PH-20, New York, NY 10032, <https://www.ohdsi.org/>.

²⁶ See the OMOP Common Data Model documentation, housed at GitHub, <https://github.com>.

²⁷ Analytical Information Markup Language (AnIML), an emerging ASTM XML standard for chemistry data sanctioned by ASTM Subcommittee E13.15. Additional information available at <https://www.animl.org/>.

standards) have also made ongoing efforts to advance the management of laboratory data, though some require a license fee/commercial approach to generate data using standard formats.

5.6.2 Blockchain-based authentication and identity management was developed in the late 2000s, though it is in its infancy in relation to laboratory and medical data systems. The technology acts as a distributed ledger that enables multiple computer servers to access, reconcile, and lend integrity to a data record. Used initially with the management of digital currency records, some in the medical and laboratory communities are proposing the use of blockchains to improve data portability and security, particularly in the scope of electronic health records (EHR) and LIS. As data integrity, sharing, and portability take on greater importance for laboratories, blockchain technology offers a solution by storing, distributing, and allowing access to records such that all participants can interact without pre-existing trust (because of cryptographic techniques) while maintaining clear audit trails.

5.6.3 End-to-end data integration in laboratory informatics is more a concept than a reality, yet years of informatics software development have seen the functionality of what was previously found in several systems make its way into a single system. The functionality, for example, of a LIS, traditionally reserved for clinical use, has found its way into what was traditionally a more research-based LIMS. This convergence of system functionality continues, with enterprise resource planning functions finding their way into laboratory informatics applications as well. Most commonly we find separate systems integrated via APIs and data transfer protocols; however, differences in data formats between systems and unclear data standards make merging laboratory and business data (to mine it effectively) more challenging.

6. Laboratory Informatics Workflow and Sample Life Cycle

6.1 *End-to-End Business Integration*—Laboratory informatics solutions are central to the broader business that it supports, as illustrated in Fig. 5. Independent drivers create demand that feed into the laboratory environment. These factors are typically industry specific and may include clinical protocols, research projects, sales, production planning, stability and environmental monitoring, and investigations. Auxiliary and concurrent workflows also exist that are driven by information flowing in or out of the active laboratory informatics workflow. Testing status, intermediate/in-process results distribution, investigation requirements, instrument calibration and maintenance, standard and reagent management, and document management (for example, SOPs, change control, and validation) are parallel processes that interface with the informatics solution. Upon completion of laboratory testing, a series of broad workflows within the larger business may be initiated that provide feedback to the internal business or external customers. Data analytics and trending, performance metrics, dispositioning of samples, and data archiving are internally focused, while certificates of analysis (COAs), sales and distribution, sample storage, and external communication distributions may be delivered externally for use by customers.

6.2 *Laboratory Informatics Workflow Introduction*—The laboratory informatics workflow model shown in Fig. 6 provides a generic representation of the workflow in a typical analytical laboratory. Workflow steps shown are aligned with the system functions described elsewhere in this guide. The purpose of the workflow diagram is to depict the laboratory

ASTM E1578-18

<https://standards.iteh.ai/catalog/standards/sist/738de0fd-5b6e-4509-9a23-f43a97a587a0/astm-e1578-18>

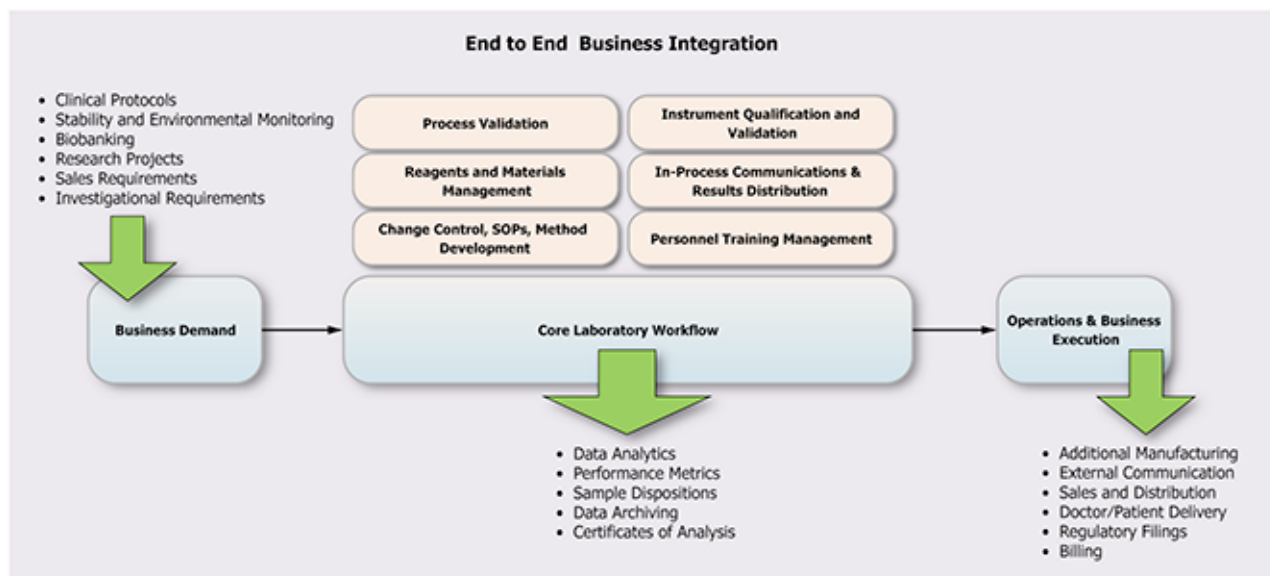


FIG. 5 End-to-End Business Integration