



**Informatics
in the
Environmental
Laboratory**



Can a laboratory operate successfully without the use of laboratory informatics? To some, the answer is always “yes, informatics makes it easier, but it isn’t a requirement”, and that is true in a technical sense, but in reality, the practical answer is “not anymore”. The use of informatics in the environmental industry is an example we’ll be looking at in this piece.

The environmental industry is a complex mix of government institutions, academic labs, and private facilities that have to address a large workload of testing and research all carried out under strict sets of regulations and guidelines that frame their operations. You could do this work in a paper-based setting, but the cost and inefficiencies of their operations would render the organization ineffective and obsolete in modern times. Today’s labs - faced with budget limitations, difficulties in finding qualified people, and a heavy workload - will find laboratory informatics as essential to operating as successful, and where appropriate, profitable, institutions.

We’ll begin by looking at the origin and current structure of the U.S. Environmental Protection Agency (EPA), the regulatory environment it manages, and then, how informatics fits within an environmental lab’s scope of operations. Most people, the author included, only think about the EPA in terms of how it affects us through things like water and air quality, but the scope of its activities is rather wide and still, after 50 years of operations, being redefined: the Supreme Court is going to hear a case in 2022 that could have a massive impact on U.S. water quality¹.

Creation and structure of the EPA

In 1969, oil slicks on the surface of Ohio’s Cuyahoga River mixed with chemicals dumped into the water and spontaneously caught fire burning down railroad bridges and causing \$100,000 in damages. That plus growing concerns about air and water quality, and oil spills in California, caused the government to create the EPA and begin reversing decades of neglect and environmentally detrimental practices. Those same concerns exist today in the form of issues like water quality in Flint Michigan and Jackson Mississippi, as well as tracing disease progression through the analysis of wastewater.

The EPA’s laboratory enterprise² consists of a set of three types of laboratories: program labs, regional labs, and research labs, all of which operate under the Office for Research and Development. Their description from the EPA website is as follows³:

“Research Laboratories have primary responsibility for providing scientific data in support of decisions by the EPA Regional Office’s environmental programs, for addressing the comprehensive needs of the Regions, and for informing immediate and near-term decisions on environmental conditions, emergency response, compliance, and enforcement.” Their work includes:

- Lab services and applied sciences divisions, and,
- Regional drinking water lab certification programs.



“National Program Laboratories have primary responsibility for conducting science to support specific programs and support decisions on regulations, compliance, and enforcement at the national level.” The program labs include a:

- Radiation lab,
- Vehicle and fuel emissions lab,
- Drinking water and wastewater lab network,
- Analytical chemistry lab (pesticides),
- Microbiology lab, and
- Environmental response lab network.

“Research and Development Laboratories conduct the research for EPA that provides the foundation for credible decision-making to safeguard human health and ecosystems from environmental pollutants.” They include centers for:

- Environmental measurement and modeling,
- Computational toxicology and exposure,
- Public Health and environmental assessment, and,
- Environmental solutions and emergency response.”

The EPA’s regulatory guidelines

Environmental testing goes beyond the governmental structure to include a host of academic and private labs carrying out water and air quality tests as well as issues such as mold testing; these may be done through walk-in/drop-off testing facilities and mail-in services. All must meet the testing requirements set up by the EPA.⁴ Among those enforced by the EPA are its good laboratory practices (GLP) guidelines which are used to enforce data quality and lab management practices. They are similar to the US Food and Drug Administration (FDA’s) guidelines and include the same issues, including validation of systems, data quality, etc.

On April 8, 1995, the EPA issued a letter⁵ “Guidelines for Study Rejection Based on GLP Considerations” that emphasizes the importance of certain aspect of the GLP guidelines. The items noted below, taken from the letter, are those where laboratory informatics can play a key role. The elements are described within three categories. Note: the text presented is extracted from the letter, though some details (e.g., references to congressional records) have been omitted. The full letter can be viewed through the link in the footnotes.⁵

Category I

“These items for consideration would lead to study rejection in almost all cases due to the seriousness of the violation(s) and the inability of the Agency to rely on the validity of the data.”

- “Evidence data are falsified”;
- “Evidence that the Final Report falsely reflects the raw data to the extent that conclusions in the submitted study could possibly be invalid”;
- “Evidence that there was no quality assurance during the conduct of a study conducted after October 16, 1989”;



- “Failure of the Quality Assurance Unit to conduct study specific inspections or maintain records of studies conducted after October 16, 1989”;
- “Evidence that test control or reference substance characterization data are falsified or confused with other test material; or evidence that substances may have been contaminated or adulterated and subsequently used in the conduct of a study”;
- “Failure to record and archive critical raw data, including study specimens/tissues; inability to reconstruct a study due to the absence of raw data that the Agency has required and/or relied upon to make a regulatory decision”; and
- “Use of unhealthy test animals or failure to maintain test animal care facilities to the extent that conclusions in the submitted study are likely to be unreliable”.

Category II

“This category includes items that would allow the individual disciplines to question the scientific conclusion made in the study. However, the registrants may be able to submit additional data which demonstrates that the violations do not affect the outcome of the study and should not be rejected. “

- “Failure to characterize or appropriately define the test, control, or reference substances”;
- “Failure to engage qualified or properly supervised individuals in the conduct of a study”;
- “Failure to adequately inspect, clean, maintain, test, calibrate and/or standardize equipment and/or maintain proper records of these activities”;
- “Failure to provide appropriate animal identification of the test system in a toxicity study allows potential confusion as to test group designation”;
- “Evidence that the testing conditions deviated grossly from the sponsor-approved protocol to the extent that conclusions in the submitted study are likely to be unreliable”; and
- “Evidence that the sponsor approved protocol was significantly inadequate or that a study contains gross, unauthorized/undocumented deviation from the sponsor approved protocol. “

Category III

“GLP items concerning the study and facility management are included in this category. These violation indications indicate a lack of control over the study (and testing facility) by the sponsor. While serious in nature, OPP may still be able to accept the study if sufficient raw data exists to reconstruct the study to ensure the validity of the data and the conclusions. These

violations should warrant careful monitoring of any future studies submitted by the registrant and/or testing facility.”

- “Gross failure to maintain Standard Operating Procedures; failure to follow laboratory SOPs without documentation in the raw data and/or written authorization from the Study Director.”

The role of laboratory informatics in environmental work

One thing that should be noted in the items above is the heavy emphasis on both documentation and cross-referencing of data and supporting information. Are people qualified? Is equipment calibrated and in good working order? Can data be traced to its source? Have standard operating procedures (SOPs) been adhered to? Trying to meet all this in a paper-based system isn't going to work, or else be prohibitively expensive and inefficient. If a five-year-old study is being audited, are you going to be able to find all relevant material in a paper-based system?

The work of environmental labs can be divided into two groups: research and service (whose emphasis is on routine, high-volume testing). The test procedures come from the EPA and other standardized sources and as a result, have a lot in common with clinical chemistry labs whose work is based on standardized testing and has greatly benefitted from the use of integrated laboratory informatics solutions. This includes systems where there is a dependence on automated instruments and communications between devices for data entry, worklist generation, and communications.

Laboratory informatics has several tiers of systems that can be applied to lab work to make it more effective and efficient. The top tier consists of electronic laboratory notebooks (ELNs) and laboratory information management systems (LIMS). ELNs act as an electronic diary of laboratory work containing project descriptions, observations, charts, data, information, etc. LIMS helps structure routine analysis workflows, keeping track of work status, results, documents, personnel qualifications, and report generation.

Supporting those are scientific data management systems (SDMS) and laboratory execution systems (LES). An SDMS would be used to keep track of files and data that don't fit an ELN's or LIMS data structures, including images, instrument-generated charts, etc. An LES manages the execution of specific laboratory procedures so that all relevant data is collected and transferred to an ELN, LIMS, or SDMS. LESs are primarily used in regulated environments.

Below that are instrument data systems (IDS) which are a combination of instruments and computer systems. Common examples are chromatography data systems (CDS) connected to one or more chromatographs, mass spectrometers, and almost any major instrument. Finally, we find devices like pH meters, balances, and other devices that have no databases associated with them and have to be programmed to be used with upper-tier systems. Their data output can be manually entered into a LIMS, ELN, or LES, but in regulated labs, the input has to be verified by a second individual.

Our primary focus will be on the top-tier systems, the LIMS and ELNs; however, all levels play a significant role in managing lab operations, improving productivity, and meeting the EPA's regulatory guidelines. Let's examine those top-tier systems within both research labs and service labs.

Research informatics

Research work is characterized by three elements: documentation, flexibility, and experiment execution. Much of the emphasis is on the documentation that describes programs, SOPs, personnel qualifications, etc. Flexibility comes in changes to the direction of work as a result of the evaluation of experimental work, again, heavy on documentation. Finally, there is the collection of experimental data consisting of manually entered results and those directly obtained from electronic instrumentation. All of these functions are supported by ELNs. A LIMS and LES may play a role as well. LIMS can provide workflow management for testing and is particularly suited for large volumes of testing. An LES can become important where the documentation of strict adherence to procedures is required. Both can feed results to the ELN, though if a LIMS is being used a better procedure would be to connect the LES to the LIMS so that all test results are in one database and then aggregated into the ELN as needed. All of this directly supports the regulatory requirements outlined above. Note, however, there is no standardized workflow model for research unless one is imposed by an institution's guidelines.

Service laboratory informatics

Service laboratories make up the bulk of organizations doing environmental work, carrying out testing in government, academic, and private labs that number in the thousands. They range in size from small facilities with a few people working in a single lab space to those with multiple buildings and a large staff. The workflow of a testing facility is consistent across industries, and, as a result, products such as LIMS can be designed and structured to meet very specific needs, building on a common base. These systems can be set up to support any sized lab and meet regulatory requirements. (Meeting regulatory requirements is something the user undertakes but LIMS vendors have tools and experience that can assist users.) A LIMS supports:

- Sample login and sample storage,
- Generation of worklists of samples to be processed,
- Entering test results of testing either manually or electronically,
- Results review and approval, and,
- Reporting the result to those who submitted the test request.

A LIMS may also be supported by an SDMS and LES as needs arise, similar to that described above for ELNs. Lab productivity can also greatly benefit from having IDSs connected to the LIMS; the LIMS can then generate worklists that are sent to the IDS and when the work is completed have the results electronically entered into the LIMS.

One aspect of environmental work is the need to collect samples in the field or have them provided by those outside the lab. They may be homeowners looking at water quality and air quality (e.g., radon testing), specialists collecting mold samples, researchers collecting wastewater samples to test for COVID-19 or other diseases, or companies providing routine samples from their facilities to meet EPA requirements.



The remote sample login feature of a LIMS can be of benefit to this service lab work. Rather than waiting to get back to the lab, the sample collection can be done in the field, with samples logged in as they are collected and barcoded labels printed and placed on the samples. This help supports chain-of-custody and data quality management efforts. Once back in the lab the sample labels can be scanned, and the sample stored and automatically listed, thus being ready for analysis.

In addition, the remote sample login and submission feature can be used to benefit relationships with high-volume sample clients. The client can, under strict access controls, log samples into the LIMS and send the actual samples to the lab to be analyzed. The client may receive a cost reduction per sample to reflect the reduced administrative workload in the lab. A lab might provide customers submitting high volumes of samples with sample containers and the ability to print barcoded labels for those containers. The reduction in workload in the lab might be reflected in the cost per sample processed. Beyond that, the client may be given access to the LIMS to check sample status and reporting, again with careful security controls in place. Given this, the LIMS should ease the administrative workload on laboratories including making it easier to generate client reports, monthly reports of a lab's performance metrics, and billing information that can be accessed directly by accounting (with security controls in place) or provided as a summary report.

In closing...

The importance of environmental labs is going to grow as water and air quality become increasingly stressed due to human activities, while the effects of climate change on all aspects of living - including pesticide use, dealing with invasive species, and water conservation – also put further downward pressure on ecosystems. Recovery from recent large storms is going to require closer controls on wastewater handling, as affected areas rebuild and have to re-engineer potable water supplies. Infrastructure changes and re-development are going to add to that as distribution systems age and have to be replaced or upgraded. This all puts more workload on environmental labs.

As such, laboratory management is going to have to re-evaluate its priorities on system design and utilization. Where the common method is to add and update information management systems from paper-based methods, to spreadsheets, and then to more sophisticated and capable lab informatics, environmental laboratories may have to skip over the first two steps and delve into LIMS and ELNs as early as possible in a labs history to provide the productivity needed to be successful.

References

¹ <https://thehill.com/policy/energy-environment/3669504-supreme-court-to-hear-case-that-could-have-massive-impact-on-us-water-quality/>

² <https://www.epa.gov/labs>

³ <https://storymaps.arcgis.com/stories/1fb957b6db6f4a39ae9c7c8387bc8493>

⁴ <https://www.epa.gov/guidance>

⁵ <https://www.epa.gov/sites/default/files/2013-09/documents/glpstudyrejection.pdf>

