Solution towards a vendor-neutral secure data transfer process between LIMS/ELN and LC-MS/MS instruments for bioanalysis

Mark E Arnold[1]; Tim Blacker[2]; Gidion De Boer[3]; Scott Davies[1]; Blair James[2]; Neil J Loftus[4]; Burkhard Schaefer[5]
[1]AAPS Data Storage Group, Arlington, VA; [2]SCIEX, Concord, ON; [3]Thermo Fisher Scientific - Digital Science Business Unit, Breda, Netherlands; [4]Shimadzu MS/BU, Manchester, United Kingdom; [5]BSSN Software, part of MilliporeSigma, Darmstadt, Germany

Overview

- Data integrity (DI) is a key consideration in a regulated bioanalytical laboratory environment to meet the needs of the ALCOA principles; the data and information must be attributable, legible, contemporaneous, original and accurate, in addition to being complete.
- One of the areas of concern relates to the bidirectional data transfers between LC–MS and information management (IM) systems (e.g., LIMS and ERP).
- The gap between the health authorities DI expectations and the availability of software that meets these expectations has resulted in a proposed solution using a vendor neutral XML file format which is digitally signed, based on the ASTM AnIML XML schema.

Introduction

Current approaches in bidirectional data transfers between LC–MS and information management (IM) systems (e.g., LIMS and ERP) are based on text file formats which require time consuming quality control steps to mitigate DI risks. This has resulted in an increased number of citings during health authority inspections highlighting the need to develop a better solution which is secure and vendor neutral.

Current challenge: common data transfer cycle using a text file format for data exchange

The data transfer cycle shown below highlights a common data transfer cycle using a text file [.txt] format as a sequence file (IM>LC-MS) and export file (LC-MS>IM). The export file includes all the results of the analysis can be saved at any location. It can be modified without any audit trail while in transit. Manual quality control checks of the data transferred (ranging from spot checks all the way to 100% check of the raw data) are therefore performed to ensure the integrity of the data.

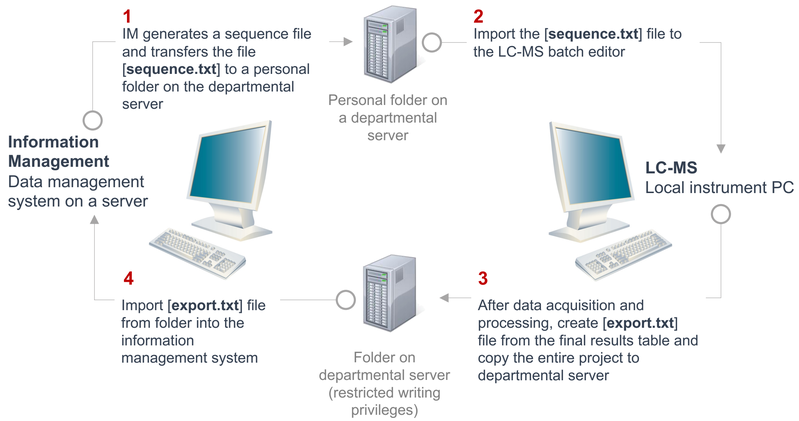

Figure 1. A common data cycle in a regulated bioanalytical workflow using .txt based data transfers, highlighting risks to data integrity and data security.

Proposed solution: a digitally signed vendor neutral xml file format for data exchange

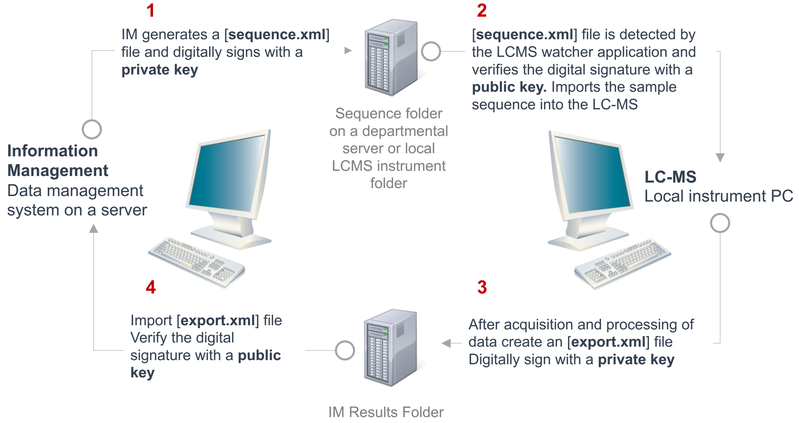

Figure 2. Proposed way forward. Replacing text file formats with digitally signed XML files. The proposed solution will have a zero impact on existing user workflows with the data file security being managed by the information management system and the LC-MS software.

Proposed solution

- Worklist/result information is transferred as a digitally signed vendor neutral XML file format
- The data model is flexible and scales beyond LC-MS workflows; open to all instrument providers; reuses existing W3C and AnIML schemas.

Improving data security by signing an XML file with digital signatures

Provided by World Wide Web Consortium (W3C) Signature Working Group;
https://www.w3.org/Signature/
https://www.w3.org/TR/xmldsig-core2/

W3C XML DSIG Characteristics

- Attributable – you can tell who made a signature
- Tamper-evident – you can tell if data has changed since signing
- Multi-signature support – multiple people and/or systems can sign a data set (approval workflows)
- Automated validation – before accepting a file, instrument or LIMS can confirm its provenance
- Integrates with enterprise PKI – some organizations already have certificate management infrastructures
- Stored in XML – easy to read and write, broad developer experience
- Royalty free, widely implemented – free libraries for Java, .net, …

Workflows and the ‘user experience’

- **Generate an IM sequence file**. The first step is the creation of a vendor neutral sequence file [**sequence.xml**]. This file includes the following fields; data file name, sample name, sample ID, tray position, vial position, injection volume, dilution factor, type (standard, QC, unknown). The [**sequence.xml**] file is then sent to folder location [**sequence folder**] set up by the system administrator.
- **Convert the IM sequence file to a LCMS instrument file**. A ‘**watcher connection**’ application running on the LC-MS automatically detects the IM [**sequence.xml**] file and verifies the digital signature. If the file has not been changed or tampered with, the file is converted into the LC-MS instrument sample list/batch file format. This process is automatic and hidden from the user. The user will simply work in the ‘usual way’ opening a sample list from a folder location using the LCMS instrument software, acquire data and review results.
  - If an **Error/Tampering/Unauthorized Change** is detected – error messages are first displayed to the user and the sequence.xml is now moved to a folder called **Error** which will exist under the Sequence Folder
  - The Audit trail will register any changes to the System configuration and LIMS configuration.
- **Review results and send to the IM**. On the LC-MS instrument data review software the user will simply click a ‘Send results to the LIMS’, the results are converted in an [**export.xml**] file and the application automatically signs the file with a private key. The [**export.xml**] file is sent to the **IM Results Folder.**
- **IM importing results from the LC-MS instrument**. In the final step of the data cycle, the IM verifies the digital signature of [**export.xml**] file and if successful stores the results in the IM system.
  - If an **Error/Tampering/Unauthorized Change** is detected – error messages are first displayed to the user and registered in an IM **error** file.

Conclusions

Data security is significantly increased

- The current data transfer cycle requires manual and time-consuming quality-control steps to mitigate the DI risks. In this proposal, the goal is to negate quality control steps and meet the need of health authorities DI expectations.

The technology is widely adopted and a vendor neutral solution

- The technology is open for all vendors to use and is intended as a generic solution.
- No bespoke interfaces are required, reducing implementation and validation cost.
- The key point is that if anyone tampers or changes the sequence or export file the change is recorded. All tampered files are detected.

It is designed to make a near-zero impact on the current user experience

- The changes are all designed to work ‘under the hood’ and will have a zero impact on current user workflows/standard operating procedures.
- In this proposal the fundamental point is the data security is managed by the software and is not exposed to the user.

References

Improving data integrity in regulated bioanalysis: proposal for a generic data transfer process for LC–MS from the European Bioanalysis Forum. Bioanalysis (2020) 12(14), 1033–1038.

Visit http://www.instrumentlink.org for more information.