Lumetics LINK^{TM} Database Export

INTRODUCTION

LINK is a software solution for automated extraction, processing and reporting of data that has been produced by analytical instruments or process equipment. LINK will search network locations for any new datasets produced by 100's of different analytical instruments/process equipment, parse all useful data from these digital files (including 'xy' curve data and instrument parameter calculations), and copy this data to a centralized LINK database residing on your network. The powerful UI allows up to 5000 distinct sample datasets to be displayed individually, statistically aggregated, and grouped based on user-specified criteria and visualized in the form of comprehensive charts/tables/images. Metadata may be automatically extracted from several different sources and utilized directly to produce dynamic fully customized Word/PDF reports. All LINK functionality is optionally 21 CFR Part11 compliant.

The LINK software is a client/server-based architecture, where the LINK client is a dedicated portable web-based application that connects to the LINK server present on your network via port 8080 (default). Most LINK users will utilize the LINK client to automate data extraction/processing/reporting and utilize the LINK database for long term storage of project data and corresponding analysis visualization preferences.

This document outlines the ability to query the LINK database and export the results to either an XLSX and/or export to direct database connection. This functionality requires a specialized license supplied by Lumetics. The LINK database Export functionality requires a minimum of one Lumetics LINK^{TM} Annual License Subscription and:
- Option 1: Export to XLSX requires a single Lumetics LINK^{TM} Export Annual License Subscription
- Option 2: Export to direct database connection requires a single Lumetics LINK^{TM} Export Annual License Subscription and a single Lumetics LINK^{TM} Staging Database Annual License Subscription

Details on both export options will be illustrated below with examples.

OPTION 1: EXPORT TO XLSX

This option is usually selected if the end user wishes to query the LINK database and export the results to an Excel XLSX. All LINK Record IDs found in the search results will be applied to measurement summary table templates from within the LINK software. This functionality requires a minimum of one Lumetics LINK^{TM} Annual License Subscription and a single Lumetics LINK^{TM} Export Annual License Subscription. There are two locations within the LINK software that the Export to XLSX feature is offered: data from within an analysis template (inclusive of analysis metadata) and data from cross-project searches establishing within the Measurement Search.

Analysis Template Export to XLSX
The measurement summary table is a flexible tabular summary that will display any LINKdb Field (Instrument or User-defined) for all the measurements isolated in the dashboard's filter panel. The measurement summary table is template-based, and once created may be utilized in any dashboard. There may be up to 10 measurement summary tables in a single dashboard. Right-click on a tabular summary within the analysis template to view the configuration and manipulation options available. Within the right-click context menu, the "Export Data" option is available. Export the data from the selected measurement summary table or all measurement summary tables in the dashboard, directly into an Excel XLSX. Optionally include the notes block on the current dashboard in the export as well. The export path must be visible to the LINK server computer.

Export Data

Please select the location from the drop-down list where the data should be exported.

[Excel Workbook ▾]

Please select the data to export:

◉ The current Measurement Summary Table Template (Default Template)

◯ All Measurement Summary Table Templates in this Dashboard

Please select the data export options:

☐ Export Notes Block on Current Dashboard

Export Path: [C:\Users\Krist\Box\Sales\2_Training Presentations] [Browse...]

File Name: [   ]

[Cancel]   [Next]

Measurement Search Export to XLSX

The measurement search features allow a user to search all Projects for measurements meeting specified search criteria. Criteria can range from imported data to raw data calculations. The search type includes numeric values, text values and specific data/time values. These measurements will then need to be linked via a common field to allow for a pool of measurements to be selected and evaluated against the specified quantitative filter expression (e.g., all measurements where Formulation ID = ABC, and both MFI + Empower results are in a specified range). And, where applicable, multiple measurements from the same Instrument will be averaged together. The default is set to LINK Record ID, where each measurement shall be evaluated independently against the specified quantitative criteria. Once measurements are found, the measurements have multiple actions. This feature can be found on the main LINK page (magnifying glass toggle).

The search criteria may be chosen from a list of options or calculated from raw data. The choose from list option allows a user to enter one or more filters to search for measurements. Filters can be selected to search for numeric or text metadata. For example, the filter can be applied to only display measurements with (e.g.) *Particle Concentrations 2-10 µm < 100* and *Melting Temperature Tm > 50 °C.* The Calculate from Raw Data option allows a user to enter one or more filters to search for measurements. Filters for this selection are raw data calculations. For example, the filter can be applied to only display measurements with (e.g.) *Particle Concentrations vs. Particle Diameter in a range of 2-10µm and acceptance criteria of < 100.*

Once the search criteria has been selected, the user may select "search for measurements". The data found in the search results may be exported into an excel workbook by selecting "Export Results".
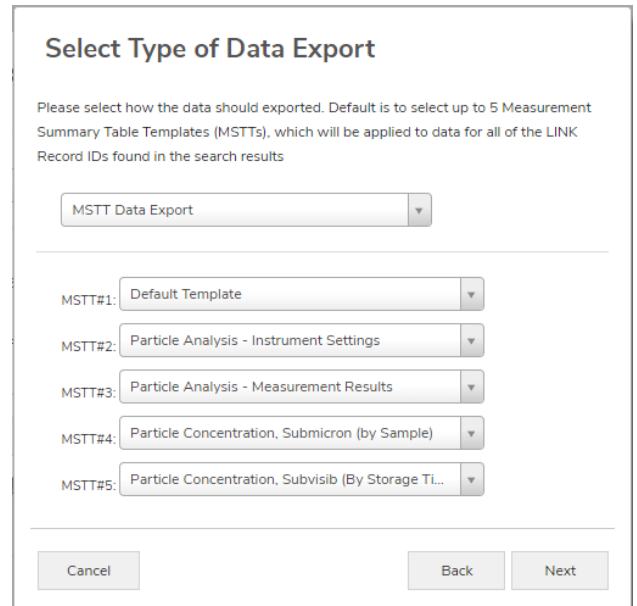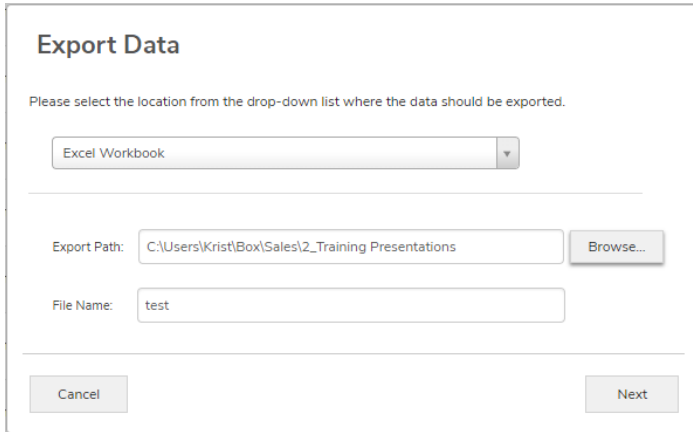
Measurement Search

Show [100 ▾] entries   Search within results: [   ]

☑	Project	Instrument Name	Import Date	Sample Name	File Name with Path
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 5.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 5.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 5.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 5.7	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 5.7	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 5.7	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 6.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 6.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His 6.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His/Arg 5.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx
☑	S217	Prometheus	2022-11-01 10:10:22.280000	His/Arg 5.2	C:\Users\Krist\Box\Customers\2_Customer Data - CONFIDENTIAL\Bayer - Germany - Julian\Bayer - Niklas\Prometheus - unsupported formats\BAY3344257_Prometheus_201005_OVERVIEW.xlsx

Showing 1 to 100 of 293 entries   Previous 1 2 3 Next

[Export Results...] [Copy Measurements Into Project...] [New Search...] [Close]

The exported results may be copied to a specified data location (accessible to the LINK server) and file name. All LINK Record IDs found in the search results may be applied to up to 5 measurement summary table templates from within the LINK software. When selecting "next" the user has the option to run the data export immediately (added to the job queue within the jobs table) or schedule the export for an hourly, daily, weekly, or one-time export selection.

DIRECT DATABASE CONNECTIVITY

This option is usually selected if the end user wishes to query the LINK database and export the results via a direct database connection. LINK will export to a database provided by Lumetics, that may be queried directly, and managed by the end user. External database exports offer the additional advantage of being queried directly for seamless ingestion to other platforms such as LIMS, ELNs, Spotfire, Tableau etc. All LINK Record IDs found in the search results will be applied measurement summary table templates from within the LINK software. This functionality requires a minimum of one Lumetics LINK^{TM} Annual License Subscription, a single Lumetics LINK^{TM} Export Annual License Subscription, and a single Lumetics LINK^{TM} Staging Database Annual License Subscription.

Analysis Template Export Direct Database Connection
The measurement summary table is a flexible tabular summary that will display any LINKdb Field (Instrument or User-defined) for all the measurements isolated in the dashboard's filter panel. The measurement summary table is template-based, and once created may be utilized in any dashboard. There may be up to 10 measurement summary tables in a single dashboard. Right-click on a tabular summary within the analysis template to view the configuration and manipulation options available. Within the right-click context menu, the "Export Data" option is available. Export the data from the selected measurement summary table or all measurement summary tables in the dashboard, directly into an established direct database connection. Optionally include the notes block on the current dashboard in the export as well.

Measurement Search Export to Direct Database Connection

The measurement search features allow a user to search all Projects for measurements meeting specified search criteria. Criteria can range from imported data to raw data calculations. The search type includes numeric values, text values and specific data/time values. These measurements will then need to be linked via a common field to allow for a pool of measurements to be selected and evaluated against the specified quantitative filter expression (e.g., all measurements where Formulation ID = ABC, and both MFI + Empower results are in a specified range). And, where applicable, multiple measurements from the same Instrument will be averaged together. The default is set to LINK Record ID, where each measurement shall be evaluated independently against the specified quantitative criteria. Once measurements are found, the measurements have multiple actions. This feature can be found on the main LINK page (magnifying glass toggle).

The search criteria may be chosen from a list of options or calculated from raw data. The choose from list option allows a user to enter one or more filters to search for measurements. Filters can be selected to search for numeric or text metadata. For example, the filter can be applied to only display measurements with (e.g.) *Particle Concentrations 2-10 μm <100* and *Melting Temperature Tm > 50 °C.* The Calculate from Raw Data option allows a user to enter one or more filters to search for measurements. Filters for this selection are raw data calculations. For example, the filter can be applied to only display measurements with (e.g.) *Particle Concentrations vs. Particle Diameter in a range of 2-10μm and acceptance criteria of < 100*.

Once the search criteria has been selected, the user may select "search for measurements". The data found in the search results may be exported to a direct database connection by selecting "Export Results".

The exported results may be copied to the specified data connection. All LINK Record IDs found in the search results may be applied to up to 5 measurement summary table templates from within the LINK software. When selecting "next" the user has the option to run the data export immediately (added to the job queue within the jobs table) or schedule the export for an hourly, daily, weekly, or one-time export selection.

DIRECT DATABASE CONNECTIVITY – STAGING DATABASE

Lumetics will charge an annual subscription fee on a per LINK server basis, to write to each individual LINKsdb. Whether an additional database in a single Mariadb database server, or multiple discrete database servers. Each database server + database name will be an optional line item in the license file, with an expiry date.

Lumetics creates the table schema automatically when Lumetics installs MariaDB as part of the Staging DB component installer. Lumetics can supply the script if the end user has their own MariaDB instance.

Each Staging Database will be able to support up to 1 million results. For simple, direct extraction of the data without aggregation, 5 million or more results are possible with basic queries. LINK is designed to prevent duplicate rows with identical data to be inserted into the database. An identical row is defined by creating a large string containing the following information:

       o   The project ID

- o   The Measurement Summary Table Template (MSTT) ID
- o   The MSTT header + value for a given row for each column in the table

This string is then used to create a "hash key" which will generate a unique ID based on the contents of the string. If this hash key already exists in the database, it means that a row containing this project/measurement summary table, and identical results, is already present, and the LINK software will not import it again.

MariaDB-based staging database for the Lumetics-supplied installer. If an existing MariaDB database is available that was not created with the Lumetics installer, then additional manual instructions will need to be applied, which is not covered in this document.

DIRECT DATABASE CONNECTIVITY – SERVER SPECIFICATIONS

The specifications of the Staging Database can be varied, depending on the loads the user expects to be placed on pushing data from the LINK Server to the Staging Database, and from other applications (LIMS, ELNs, Spotfire, etc) pulling data from the Staging Database. For moderate usage, starting with the specifications below should be adequate, and can be increased as loading/usage increases.

Effective CPUs	4
RAM	8+ GB
Hard Drive Free Space	≥ 250 GB
Operating System	Minimum: Windows 7+ or Win Server 2008+, 64-bit OS
CPU Clock Speed	≥ 4 GHz
Disk R/W Speed	≥ 3000 MB/s (typically SSD)
Network Speed	≥ 1 Gbps (between data source and LINK Server)

MARIADB STAGING DATABSE INSTALLER

Lumetics will install the MariaDB database using the LINK Staging Database (MariaDB) Installer, by following the below steps:

Run the Installer
Download and run the "*LINKdb Staging Database MariaDB Installer vX.exe*" installer on VM that was created for this purpose. This cannot be the same VM as the LINK Server installation. Running this installer will perform the following operations:
- Create a root user with a default password of *STAGEadmin@#*
- Create a database called *lumetics_staging*, which contains two tables: *staged_data* and *column_names*

Create Account
Access to the staging DB account will be established to read data and delete old data. The below commands will a) create a user, b) grant permissions to access certain tables, as well as specify what other computers they can connect from, and c) "Flush Privileges" (which basically activates the commands)

```
CREATE USER 'sdb_user'@'%' IDENTIFIED BY 'STAGEuser!@';

GRANT SELECT, DELETE, RELOAD, REFERENCES, INDEX, EXECUTE ON *.* TO
'sdb_user'@'%'  IDENTIFIED BY 'STAGEuser!@' WITH GRANT OPTION;

FLUSH PRIVILEGES;
```

For the "GRANT" command, it is set to log in from anywhere by virtue of specifying the "%" after the user name (e.g.) 'sdb_user'@'%'. If only allowing login of this user from a specific computer (e.g. a Spotfire server), then this will change to be a specific computer name in which the connection is from (e.g. 'sdb_user'@'my_spotfire_server_name_here' (replace "my_spotfire_server_name_here" with the actual name of the Spotfire server).

Listing the available columns – to see a list of all the available field names that have been exported to the staging database, simply use the following basic command:

```
SELECT `Name` FROM `column_names` ORDER BY `Name` ASC;
```

Basic Dynamic Column Example

Below is an example query a user might run manually using the MariaDB dynamic columns:

```
SELECT `Id`, COLUMN_GET(`Data`, 'Sample Name' AS CHAR),
COLUMN_GET(`Data`, 'LINK User' AS CHAR),
COLUMN_GET(`Data`, 'Export Date' AS CHAR),
COLUMN_GET(`Data`, 'MSTT Name' AS CHAR),
FROM staged_data;
```

A user can select any column names they wish to query from the 'column_names' table, and write a query based on the above syntax.

Basic Dynamic Column Example with WHERE clause

To limit the results to a specific export date (or later), an example query might be:

```
SELECT `Id`, COLUMN_GET(`Data`, 'Sample Name' AS CHAR) AS 'Sample Name',
COLUMN_GET(`Data`, 'LINK User' AS CHAR) AS 'LINK User',
COLUMN_GET(`Data`, 'Export Date' AS CHAR) AS 'Export Date',
COLUMN_GET(`Data`, 'MSTT Name' AS CHAR) AS 'MSTT Name'
FROM staged_data
WHERE COLUMN_GET(`Data`, 'Export Date' AS CHAR) > '2023-10-01';
```

Note that the WHERE clause requires the dynamic column query as well. A result of this query is presented below as an example:

```
 8 SELECT `Id`, COLUMN_GET(`Data`, 'Sample Name' AS CHAR) AS 'Sample Name',
 9 COLUMN_GET(`Data`, 'LINK User' AS CHAR) AS 'LINK User',
10 COLUMN_GET(`Data`, 'Export Date' AS CHAR) AS 'Export Date',
11 COLUMN_GET(`Data`, 'MSTT Name' AS CHAR) AS 'MSTT Name'
12 FROM staged_data
13 WHERE COLUMN_GET(`Data`, 'Export Date' AS CHAR) > '2023-10-01';
14
```

staged_data (73,028r × 5c)

Id	Sample Name	LINK User	Export Date	MSTT Name
34,684	Sample 1	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,685	Sample 2	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,686	Sample 3	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,687	Sample 4	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,688	Sample 5	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,689	Sample 6	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,690	Sample 7	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,691	Sample 8	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,692	Sample 9	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,693	Sample 10	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,694	Sample 11	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template
34,695	Sample 12	LINK Admin	2023-11-06T16:11:36.255-0500	Default Template

Duplicate Insertions

The LINK software is designed to prevent duplicate rows with identical data from being inserted into the database. An identical row is defined by creating a large string containing the following information:

(1) The project ID
(2) The Measurement Summary Table Template (MSTT) ID
(3) The MSTT header + value for a given row for each column in the table

This string is then used to create a "hash key" which will generate a unique ID based on the contents of the string. If this hash key already exists in the database, it means that a row containing this project/measurement summary table, and identical results, is already present, and the LINK software will not import it again.

Here is the basic schema for the database and tables. Change the name of the database (e.g. *lumetics_staging*) below when running in SQLYog to add multiple staging databases with different names to the same server.

```
CREATE DATABASE IF NOT EXISTS `lumetics_staging` CHARACTER SET 'utf8';

USE `lumetics_staging`;

DROP TABLE IF EXISTS `staged_data`;

CREATE TABLE `staged_data` (
  `Id` INT(11) NOT NULL AUTO_INCREMENT,
  `Hashkey` VARCHAR(255),
  `Data` BLOB,
  PRIMARY KEY (`Id`),
  UNIQUE KEY `Index2` (`Hashkey`)
) ENGINE=INNODB AUTO_INCREMENT=1 DEFAULT CHARSET=utf8 COLLATE=utf8_general_ci;

DROP TABLE IF EXISTS `column_names`;

CREATE TABLE `column_names` (
  `Id` INT(11) NOT NULL AUTO_INCREMENT,
  `Name` VARCHAR(255),
  PRIMARY KEY (`Id`),
  UNIQUE KEY `Index2` (`Name`)
) ENGINE=INNODB AUTO_INCREMENT=1 DEFAULT CHARSET=utf8 COLLATE=utf8_general_ci;
```

Deleting Data

As MariaDB is based on MySQL, standard syntax based on MySQL is used to manipulate data. For instance, to delete rows based on ID (from the example above), one might use the following query:

```
DELETE FROM staged_data WHERE `Id` > 34684;
```

To use Dynamic Columns, the query might look like:

```
DELETE FROM staged_data
WHERE COLUMN_GET(`Data`, 'Sample Name' AS CHAR) = 'Sample 1';
```

As with all delete operations, it is best to test the query with a "SELECT" statement first using the same WHERE clause, to ensure that the data being captured from the delete is correct!

CONTACT LUMETICS

For direct assistance, please contact Lumetics LINK™ Support:

E-mail: support@lumetics.com
Phone: 1.613.417.1839
Website: http://lumetics.com/