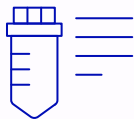# Best Practices: Data Ingestion

Benchling

Data Ingestion Methods
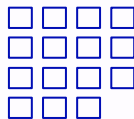
Most manual, lowest throughput

Least manual, highest throughput
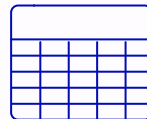
Single-Entity Import

One-Off import of entities and containers within the Benchling User Interface

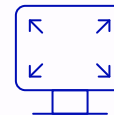## Structured Table Import

Document bulk structured data import within the notebook

Spreadsheet Import

Import Data in Spreadsheets to perform bulk actions

API Import

Leverage Benchling's API to automate import of a constant stream of data

*Benchling's Lab Automation tools are another option for ingesting high-throughput or instrument data. Given the additional licensing required, we will not include descriptions of this in this resource.

Single-Entity Import

Overview:
Using the Benchling user interface, you can create single entities. Structured data is presented as a form to be filled out, with options for placing entities within specific projects. It can be initiated through the Global Create, Registry Create, or Inventory Create buttons.

Limitations:
Single Entity import is slow for processes that will involve more than 5-6 entities; registration tables become faster for larger imports.

Resources:

- [Create and register a single entity](#)
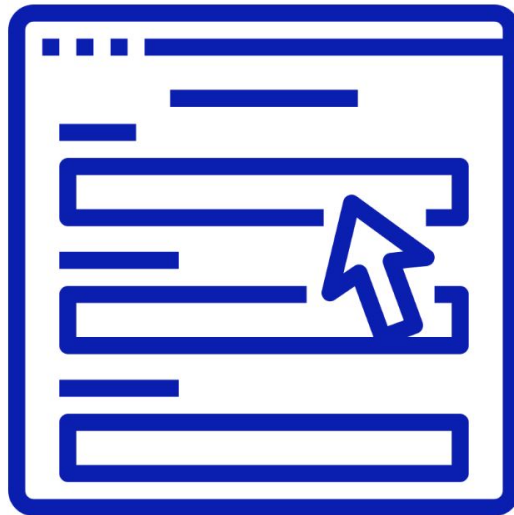
Customer Success Recommendation:

Best for import of 1-20 entities
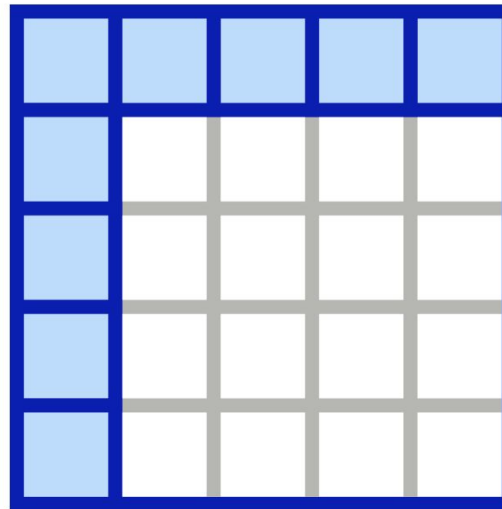
Structured Table Import

Overview:
Structured tables in the Notebook can be used to register entities **or** create, fill, and update containers. From a structured table, you can choose to fill rows manually, or harness features like worklists or inserting from spreadsheets. This method also allows you to document the context in which entities are registered by writing descriptions or procedures in the Notebook entry.

Limitations:
Large or numerous structured tables in an entry can slow down loading and performance. It is not recommended to have more than 500 rows in a structured table. Structured Table Limits

Resources:
Registering entities using Registration tables

Customer Success Recommendation:

Best for import of 20–500 entities

Spreadsheet Import

Overview:
Spreadsheet imports help ingest large amounts of bulk data. Access this from any Global Create, Registry Create, or Inventory Create, and select the spreadsheet import option.

Limitations:
You can only import only **one Schema type at a time, with a limit around 4K** . Fields within the spreadsheet must be mapped to metadata fields configured within the entity schema.

Resources:
- [Upload inventory information using spreadsheets](#)
- [Import Registry data in bulk](#)

Customer Success Recommendation:

Best for import of 500-4k entities

API Import

Overview:
Using the Benchling SDK, API, and Apps platform, programmatically import large volumes of data. Building on this allows for easily repeatable imports of data without manual steps.

Limitations:
Using the API requires technical knowledge and programming abilities—this is best done by data scientists who understand the scientific and programmatic constraints. The API also has Rate Limits for how quickly data can be ingested, however, these programs can run continuously.

Resources:
- [Developer Platform Overview](#)
- [API Reference](#)
- [Registering Entities](#)
- [Registering Entities via the API](#)
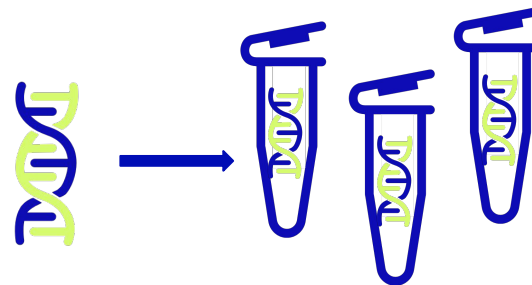- [Rate Limits](#)
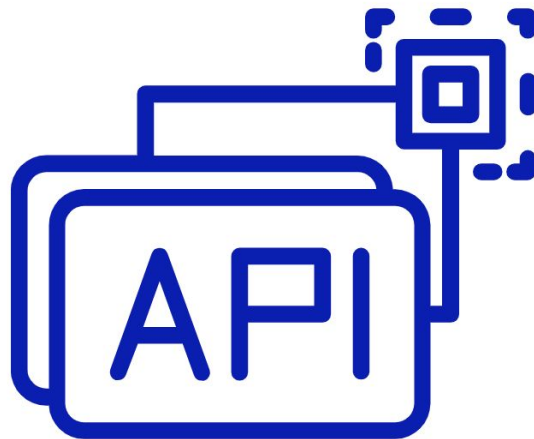
Customer Success Recommendation:

Best for ongoing & large volume entity imports (500+)

Ingesting Physical Sample Data

In Benchling, Entities and Containers are
unique data objects. Entities are contained within one or
more containers.

2 Upload Process:	Register **Entities**, create **Containers** using Entity IDs to create connections within a Notebook entry or via the UI. [Help Article]	Best for **<100 containers**, limited to one container type
1 Upload Process:	Append **Container** data to **Entity** registration spreadsheet, and upload both simultaneously. See next slide for more guidance [See next slides]	Best Suited for **500-4K containers**, limited to one container type
Programmable Process:	Leverage **Entity** and **Container** Endpoints in the API	Best for regular **>4K inputs or high-throughput processes.** Requires programming knowledge

Single Upload Process – Step 1
Register and Inventory Samples Simultaneously

When Importing and Inventorying Samples at once:

1. Go to Import entities, and after selecting your schema, check the box indicating that you would like to "transfer entities into new or existing containers"
2. Select the type of container you are adding the entities to

Note: You can only add one type of Entity Schema and one type of Container Schema at once.

Single Upload Process – Step 2
Register and Inventory Samples Simultaneously

Spreadsheet Structure for a 1 Step Process
You will need to append Container Metadata, such as which box/location they will exist in, the positions they will go, and any volumes/concentrations

Note: For multiple containers of the same entity, volume, and concentration, you can upload all simultaneously with one line, by using a comma separated list of positions.
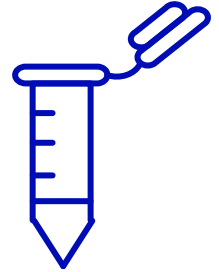
Entity Metadata — Container Metadata

	A	B	C	D	E	F	G	H	I	J
1	**Entity name**	**Metadata 1**	**Metadata 2**	**Metadata 3**	**Location Barco**	**Location Position**	**Volume**	**Volume Units**	**Concentration**	**Concentration Units**
2	Chain 1	Value 1	Value 11	Value 111	FBOX003	1,2,3,4,5	100	ul	90	mM
3	Chain 3	Value 3	Value 33	Value 333	FBOX003	6,7,8,9,10	150	ul	90	mM
4	Chain 4	Value 4	Value 44	Value 444	FBOX003	11	200	uL	90	mM

Order of Operations – Physical Sample Data

Use the principals below for any data ingestion methods

1 Determine which metadata should be captured under the **entity**, **entity lot**, and **container**

2 Take into account the *entire* process of registering entities and filling containers when deciding how best to import your sample data

3 Create and register entities prior to creating the containers (unless using the one upload process). Avoid having empty containers in your Benchling environment.

Resources for Data Ingestion

[Create and register a single entity](#)

[Upload inventory information using spreadsheets](#)

[Import Data Registry in Bulk](#)

[API Overview and Tutorial](#)