

Bulk Initial Data Loading

- [Background](#)
- [Considerations](#)
 - [Dataset Size](#)
 - [Data Availability](#)
 - [Linkage With Upstream Data Sources](#)

Background

Deployers with existing legacy data sets will need to consider how to prepopulate those records into Registry. There are several possible approaches, each with different considerations:

1. [Manual Enrollment](#)
2. [API Enrollment](#)
3. [Organizational Identity Source \(OIS\) Import](#)
4. [Registry BulkLoad Shell](#)

Considerations

Dataset Size

- For large datasets (> ~50k), BulkLoad offers the best performance.
- For small datasets (< ~100), Manual Enrollment offers the lowest setup overhead.
- In the middle, API Enrollment or OIS Import are probably the best options.

Data Availability

- All options *except* OIS Import require data to be available locally for import.
- OIS Import can connect to upstream datasets and synchronize them without requiring local copies.

Linkage With Upstream Data Sources

Organizational Identity Sources allow for records to be synced from upstream sources. After the initial load, internal database keys must be in place for [subsequent updates](#) to happen automatically.

- OIS Import will set the internal keys automatically.
- BulkLoad will set the internal keys automatically when the inbound file is configured appropriately. See the [documentation](#) for details.
- Manual Enrollment will set the internal keys if [Enrollment Sources](#) are used.
- Setting the keys via the API is not currently supported.