## 2022-05 Bulk Data

- **Short Description**
  Participants in this track will test the FHIR Bulk Data Access Implementation Guide (IG) with a focus on new and updated implementations, and features added in STU2. Participants may also begin prototyping and provide feedback on the early stage Bulk Import IG.

- **Long Description**
  Healthcare and payor organizations have many reasons to transmit data on large populations of patients, such as moving clinical data into an analytic data warehouse, sharing data between organizations, or submitting data to regulatory agencies. Today, bulk export is often accomplished with proprietary pipelines, and data transfer operations often involve an engineering and field mapping project. The bulk data implementation guide (IG) is an exciting effort by HL7, Argonaut, and SMART to bring the FHIR standard to bear on these challenges of bulk-data export. This track provides a forum for server implementors (providing bulk data) and client implementors (retrieving bulk data) to test the current IG and the updates in version 2. Version 2 incorporates learnings from implementer experience with v1.0 and has been developed over the past two years through a series of open community meetings coordinated by the Argonaut FHIR Accelerator.

  Track participants may also begin prototyping the early stage Bulk Import IG. The Bulk Import IG is intended to address use cases where an organization needs to share a large FHIR dataset with another organization or move it between systems within an organization on a schedule defined by the data provider. Inter-organizational examples include submitting FHIR data to a disease specific registry, sending information to a public health institution, or transmitting clinical information to a payor for a quality based payment program. Within a single organization, the operation can be used to coordinate ETL tasks, for example, when a client application wishes to instruct a FHIR server to load a new dataset from a local file server or a cloud storage bucket such as AWS S3.

- **Type**
  Test of an Implementation Guide

- **Submitting Work Group/Project/Accelerator/Affiliate/Implementer Group**
  FHIR-I, Argonaut

- **Track Lead(s)**
  Dan Gottlieb, Jamie Jones, Josh Mandel

- **Track Lead Email(s)**
  dani@centralsquaresolutions.com; james.jones@chip.org; jmandel@gmail.com

- **FHIR Version**
  FHIR R4 / Any Version
| Specification(s) this track uses | Bulk Data Access IG (Export): [https://hl7.org/fhir/uv/bulkdata/](https://hl7.org/fhir/uv/bulkdata/)
| Artifacts of focus |  |
| Expected participants | Sign up Sheet / Export Servers and Clients / Import Providers and Consumers |
| Zulip stream | [https://chat.fhir.org/#narrow/stream/179250-bulk-data](https://chat.fhir.org/#narrow/stream/179250-bulk-data) |
| Track Kick Off Call |  |
If you're not familiar with the FHIR Bulk Data Access IG, please review this overview of Bulk Data presentation before the Connectathon.

A track orientation session will be held on Thursday, April 28th at 12pm ET: Zoom Link

---

System Roles:

1. **Bulk Data Provider** - may consist of:
   a. **FHIR Authorization Server** (scenarios 2 and 3 only) - server that issues access tokens in response to valid token requests from Bulk Export Client.
   b. **FHIR Resource Server** - server that accepts kick-off request and provides job status and completion manifest.
   c. **Output File Server** - server that returns FHIR Bulk Data files and attachments in response to urls in the completion manifest. This may be built into the FHIR Server, or may be independently hosted.
   d. **Bulk Import Client** (scenario 3 only) - application or server that sends a ping to the Bulk Data Consumer to indicate that a FHIR Bulk Data dataset is available for export and processes status information once the import is complete.

2. **Bulk Export Client** (scenarios 1 and 2 only) - system that requests and receives access tokens and Bulk Data files.

3. **Bulk Data Consumer** (scenario 3 only) - system that accepts a notification that a FHIR Bulk Data dataset is available, requests an export of the dataset, ingests the data, and provides status information back to the Bulk Data Provider.

Scenario 1: Bulk data export with retrieval of referenced files on an open endpoint

Bulk Data Provider and Bulk Export Client follow the flow outlined in the Bulk Data Export IG to generate and retrieve a dataset.

Scenario 2: Bulk data export with retrieval of referenced files on a protected endpoint

1. Bulk Export Client registers with Data Provider, per SMART Backend Services Authorization
2. Bulk Export Client obtains an access token, per SMART Backend Services Authorization
3. Bulk Export Client follows the flow outlined in the Bulk Data Export IG to generate and retrieve a dataset, passing in the access token.
4. Bulk Export Client attempts to retrieve a dataset using the flow outlined in the Bulk Data Export IG without providing a valid access token and fails

Scenario 3: Bulk data import with retrieval of referenced files on a protected endpoint

Bulk Data Provider and Bulk Data Consumer follow the flow outlined in the Bulk Data Import IG to generate and retrieve a dataset

**Security and Privacy Considerations**

SMART Backend Services Authorization will be required to participate in Scenario 2 and Scenario 3