ClinGenDB – HTTP REST APIs

ALL functionality for retrieving, changing, deleting, and searching exposed via a *RESTful HTTP APIs*

Programming language independent
- Use your language's HTTP library

Session-less
- No cookies, no SOAP/CORBA/etc overhead

Resource-oriented, not method oriented

Uses JSON as the principal format for getting & putting data

General information about the Genboree HTTP REST API:
- **http://www.genboree.org/java-bin/showHelp.jsp?topic=restAPIOverview**
- (And linked pages containing further detail)

HTTP-Based APIs – Three Key Aspects

Three Key Aspects of HTTP-Based APIs to Consider:

1. HTTP URL
- i.e. S*ubject* of the operation; the *resource* or *entity* the request is about
- Specifies the <u>location</u> of the resource on the Internet

3. HTTP Method
- i.e. What you want to <u>do</u> to that resource
- HTTP defines a limited & sufficient set of methods:
  - **GET, PUT, DELETE, HEAD, (OPTIONS),(POST)**

5. HTTP Payload
- A <u>representation</u> of the resource or part of the resource
- *Response* payload (or "*body*") → data *from* the server
- Request payload → data *to* the server
- It will deliver JSON and will expect to be given JSON
- It will be wrapped in our ***standard envelope***

ClinGenDB – HTTP REST APIs

Using the APIs to insert, retrieve and update the following:

- The *document* for each variant, phenotype, etc.

- The *list of document identifiers* matching the search string (initially for document names only)

- The *model* for the selected *document collection*

- The *list of document collections* present in the Genboree KB (ClinGen-related prefixed by "CG")

Components of the Genboree HTTP API URL:

Resource URI	Auth. Parameters
http://{genbHost}/{rsrcPath}?{rsrcParams}	{authParams}

The URLs have a general form, with several distinct parts:

- The core of the resource URL is the **{rsrcURI}**, which is everything but the authorization parameters:
  - **{genbHost}** is the host where Genboree is running
    (e.g. www.genboree.org, genboree.cbrc.jp, etc.)
  - **{rsrcPath}** is the path identifying the resource at that host
    (e.g. /REST/v1/grp/{grpName}/kb/{kbName}/colls)
  - **{rsrcParams}** are any optional parameters
    - Generally to modify defaults, apply filters, ask for more detail, etc

- Appended to the **{rsrcURI}** are 3 required parameters Genboree uses to *authorize* and *authenticate* the user accessing/modifying the resource
  - Recall that data access is subject to ACL
  - See the REST API - Authorization page for more detail

ClinGenDB – Selected API Resource Paths

The following resource paths are some of those already implemented or which will be available in the first release:

- Each GenboreeKB is within a Genboree group ; thus all resource paths have this common prefix:
  - `/REST/v1/grp/{grp}`

How do I get the list of Collections in a particular GenboreeKb?

`…/kb/{kbName}/colls`

    ← Returns a list of collection names

The data Model for a particular Collection?

`…/kb/{kbName}/coll/{collName}/model`

    ← Returns the model/schema doc

ClinGenDB – API Resource Paths

All Records within a Collection?

`…/kb/{kbName}/coll/{collName}/docs`
   ← **Returns a list of all doc names/identifiers**

`…/kb/{kbName}/coll/{collName}/docs?detailed=true`
   ← **Returns a list of all the docs (!)**

A specific Record within a Collection:

`…/kb/{kbName}/coll/{collName}/doc/{docName}`
   ← **Returns the doc**

Basic select by property value:
- `{propPath}` is the path to the property using "." as delimiter

`…/docs?matchProp={propPath}&matchValue={value}`
`…/docs?matchProp={propPath}&matchValue={valPrefix}&matchMode=prefix`
   ← **Returns a list of doc names/identifiers**

(adding `detailed=true` would return a list of the docs themselves)

ClinGenDB – Representations in payloads

HTTP Request and Response bodies contain <u>representations</u> of the data resource stored [or to be stored] on the server

The Genbore HTTP API wraps resource representations a *standard envelope* containing two fields:

- "`data`" → value is the data structure representing the resource
- "`status`" → value contains the HTTP status code text and any extra status info

Representation of data entity available via "`data`"

```
{
  "data" : <jsonRepresentation>, /* Data entity representation (defined elsewhere) */
  "status" :
  {
    "statusCode" : "<value>",    /* HTTP response status text (see previous table) */
    "msg" : "<value>"            /* Extra info, esp. for errors. */
  }
}
```

Status information available via "`status`"
(Useful when an error is communicated, especially during development.)

When **PUT** is used to deliver such structured data,
the <u>"status" property may be omitted</u>. If present, it will be ignored.

ClinGenDB – Authentication & Authorization

Consider:
- Need to know if one is allowed to access or change resources within a GenboreeKB (as dictated by one's Role in the Group)
- RESTful HTTP design means no sessions

Thus, <u>ALL</u> requests to private (non-publicly exposed) GenboreeKBs require 3 authentication-related tokens at the end of the URL.

`&gbLogin=api_demo&gbTime=1284128280&gbToken=499a7bcc1833b56b424eb78cc3575e22f93fe858`

The REST API requires these to be appended (in any order) to the <u>end</u> of resource URLs:

1. `gbLogin`
   - Login name of registered Genboree user with sufficient access to perform method on resource.
3. `gbTime`
   - Current POSIX time (a.k.a. epoch time or UNX time) as an integer.
5. `gbToken`
   - A *single-use*, *location-* and *time-constrained* string
   - Results from a nested digest (see formula below)

`gbToken` **Formula:**

```
token = SHA1( {rsrcURI} + SHA1({gbLogin} + {userPassword}) + {gbTime} )
```