**HFCuS**

HaplotypeFrequency

Curation Service

# Input file

ASCII File, XML

## Structure

|  |  |  |  |
| --- | --- | --- | --- |
| Mandatory | XML Tag | Values / Kind of Val. | Description |
| Yes | POP\_ID | URI from POP DB | Population Identifier |
| Yes | HT | HT\_List  HT\_Lic  HT\_Res |  |
| Yes | HT\_List | List of HT |  |
| Yes | HTL | Pair of HTL\_Name, HTL\_Freq |  |
| Yes | HTL\_Name | GL String |  |
| Yes | HTL\_Freq | 0 < f <= 1 |  |
| Yes | HT\_Lic | License ID | License under which HF data is available |
| Yes | HT\_Res | G, P, gNMDP, gDKMS, n-Field, Serology |  |
| No | GT | GT\_List,  COHORT\_ID | As an alternative to GT List, NOT THE SAME IDs as used in the HFCeS |
| No | GT\_List | List of GTL | List of Genotypes |
| No | GTL | Record of GTL\_Name, GTL\_M\_List |  |
| No | GTL\_Name | GL String | Raw Data |
| No | GTL\_M\_List | List of GTLP\_Meth | Additional Information on GT, Typing |
| No | GTLP\_Meth | Pair of GTLP\_M\_Data,  GTLP\_M\_Value |  |
| No | GTLP\_M\_TYPE | String | Free Text / Predefined Tag |
| No | GTLP\_M\_VALUE | Free |  |
| No | GT\_Lic | License ID | License under which GT data is available |
| ~~No~~ | ~~METHOD\_LIST~~ | ~~List of METHOD~~ |  |
| No | METHOD | Record of  HH2016  ~~METH\_Type, METH\_Value,~~  ~~METH\_CLASS~~ |  |
| No | HH2016 | METH\_Type  METH\_Value |  |
| No | METH\_Type | String | ~~Free Text~~ Tags from List |
| ~~No~~ | ~~METH\_Value~~ | ~~String~~ | ~~Free Text~~ |
| ~~No~~ | ~~METH\_CLASS~~ | ~~String~~ | ~~From predefined list / Text~~ |
| No | METHOD\_ID |  | As an alternative to Method List, NOT THE SAME IDs as used in the HFCeS |
| ~~No~~ | ~~QUALITY\_LIST~~ | ~~List of QUALITY~~ |  |
| No | QUALITY | Record of  HH2016  ~~QUAL\_TYPE, QUAL\_VALUE,~~  ~~QUAL\_CLASS~~ |  |
| No | HH2016 | QUAL\_TYPE |  |
| No | QUAL\_TYPE | String | ~~Free Text~~ Tags from List |
| ~~No~~ | ~~QUAL\_VALUE~~ | ~~String~~ | ~~Free Text~~ |
| No | QUAL\_CLASS | String | Predefined list / Free Text |
| ~~No~~ | ~~LABEL\_LIST~~ | ~~List of LABEL~~ |  |
| No | LABEL | HH2016  ~~Pair of LABEL\_TYPE, LABEL \_VALUE~~ |  |
| No | HH2016 | Label\_Type |  |
| No | LABEL \_TYPE | Free Text | Predefined List ~~and Free Text~~ |
| ~~No~~ | ~~LABEL \_VALUE~~ | ~~Free Text~~ |  |
| No | LABEL\_CLASS | Free Text | Predefined List and Free Text |
| No | ACL | ToBeDefined  <Defaults to public/private> ?? | Access Control List |
| ~~No~~ | ~~COHORT\_ID~~ |  | ~~As an alternative to GT List, NOT THE SAME IDs as used in the HFCeS~~ |
| ~~No~~ | ~~METHOD\_ID~~ |  | ~~As an alternative to Method List, NOT THE SAME IDs as used in the HFCeS~~ |
|  |  |  |  |

## Additional Stored Values

* Timestamp
* Submitting UserID

## Direct Output/Feedback

* URI to dataset
* Method\_ID
* Cohort\_ID
* HF\_ID

# Internal Data Structure of HFCuS

## Basic Fields

The basic data structure of the HFCuS mimics the input file

## Additional Fields

To allow for curation of submitted data sets, comments can be used. They are an independent submission to the HFCuS but refer to an existing HF data set. The field COM\_REF\_SPEC can be used to specifically address a comment to a certain piece of data in the original set.

|  |  |  |  |
| --- | --- | --- | --- |
| Madatory | Data\_Field | Content | Description |
|  | COMMENT | Record of COM\_TIME, COM\_USER, COM\_REF~~\_HF~~,  COM\_LIST | Mandatory fields:  COM\_TIME, COM\_USER, COM\_REF~~\_HF~~, COM\_LIST |
|  | COM\_TIME | Timestamp | Time of addition of the comment |
|  | COM\_USER | UserID | The user of HFCuS adding the comment |
|  | COM\_REF | Pair of COM\_REF\_TARGET, COM\_REF\_ID | The HF set the comments refer to |
|  | COM\_REF\_TARGET | Free Text/COHORT\_ID, METHOD\_ID, HF\_ID, POP\_ID, ~~COMMENT~~ |  |
|  | COM\_REF\_ID | ID | Appropriate Of the above target |
|  | COM\_LIST | List of COM\_REC |  |
|  | COM\_REC | Record of COM\_TEXT, COM\_REF\_SPEC |  |
|  | COM\_TEXT | Free Text | The comment |
|  | COM\_REF\_SPEC | Free Text | Some hints what the comment is referring to |

Also, if the genotype list is available, GTs can be downloaded, HF resubmitted for the same Cohort\_ID with a different (better!) methodology.