



IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

ImmPort Data Submission Process

The ImmPort system was developed under the Bioinformatics Integration Support Contract (BISC) by the Northrop Grumman Health Solutions team for the National Institute of Allergy and Infectious Diseases, Division of Allergy, Immunology, and Transplantation. This portal is used to archive, exchange and analyze scientific data for the life science community.

ImmPort stands for **I**mmunology Database and Analysis **P**ortal.

This tutorial will describe the process for sending data to ImmPort. Please contact the helpdesk at helpdesk@immport.org before you start your data submission process. We want to help you.

The main steps in data submission to ImmPort are:

- Download the file templates for providing experiment information and experiment results
- Fill out the templates. You can use the Spreadsheet or XML. This tutorial will cover only the spreadsheet templates. If you want to use XML, please contact us.
- Save each completed template as tab-delimited text file and prepare all accompanying files
- Zip all the files together at ImmPort website to create the submission package. The combination of metadata files and the results files comprise the ImmPort Data Submission Package
- Upload the submission package and get a submission ticket number to track your package in ImmPort
- If the package is larger than 1 GB, contact us.

The Immunology Database and Analysis Portal or ImmPort is used to archive, exchange and analyze scientific data for the life science community. Before you submit data into ImmPort, we strongly recommend that you call or email us.

This tutorial will describe the process of sending data into ImmPort.

Log in to ImmPort

Download and Fill the metadata templates

Save all the metadata templates as tab-delimited files

Zip the metadata and results files in one package

Check for any issues using the Data Package Validator. If no errors exist, upload your package into ImmPort

If you experience any problems contact the helpdesk at helpdesk@immport.org. We will be happy to go through the process with you.



IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

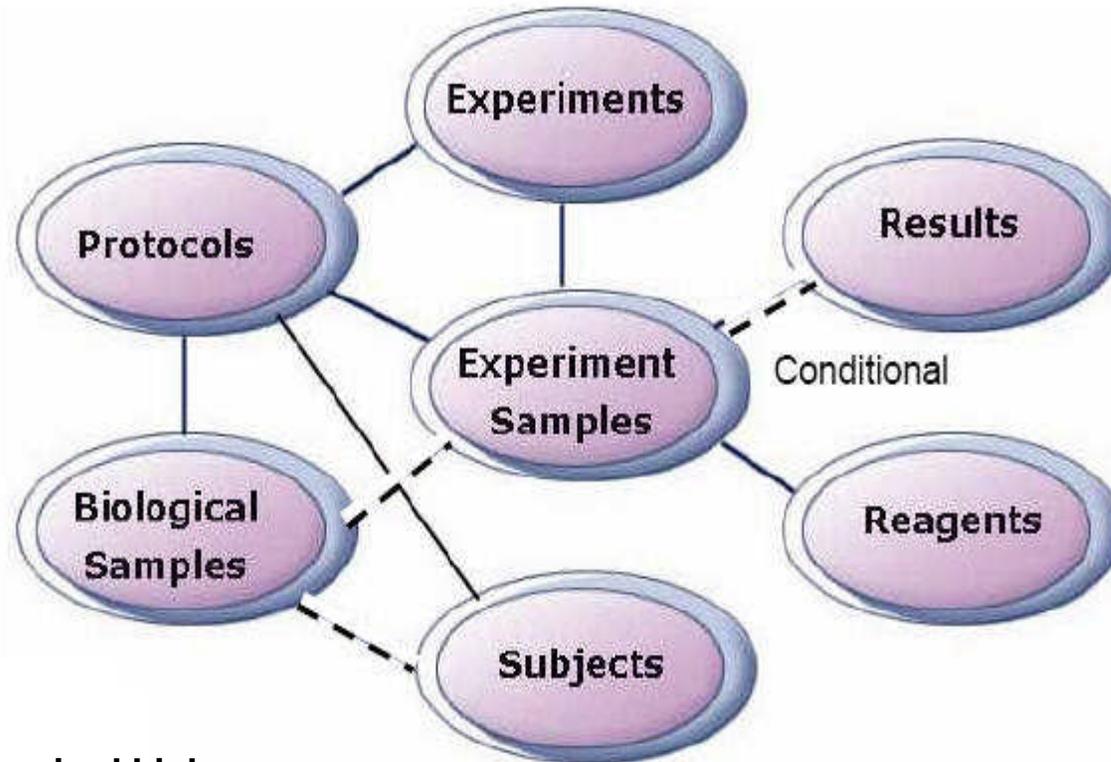
ImmPort Data Submission Process

#1: LOG INTO IMMPORT

1. Log into ImmPort at <https://www.immport.org/immportWeb/home/home.do>
At the Log on page, WebEx materials are available to the ImmPort user.
2. This page indicates some useful information available to the ImmPort user. The data submission schematic, user guide, tutorial, example packages, data package validator and XML Resources (for XML template users). The release documentation notes are also available by clicking Resources.
3. Select Submit Data from the left panel or from the Submit Data module.
4. Make sure you update your templates before using them.
5. We strongly recommend you familiarize yourself with ImmPort resources. If you have any questions regarding any of the ImmPort resources, please contact us.



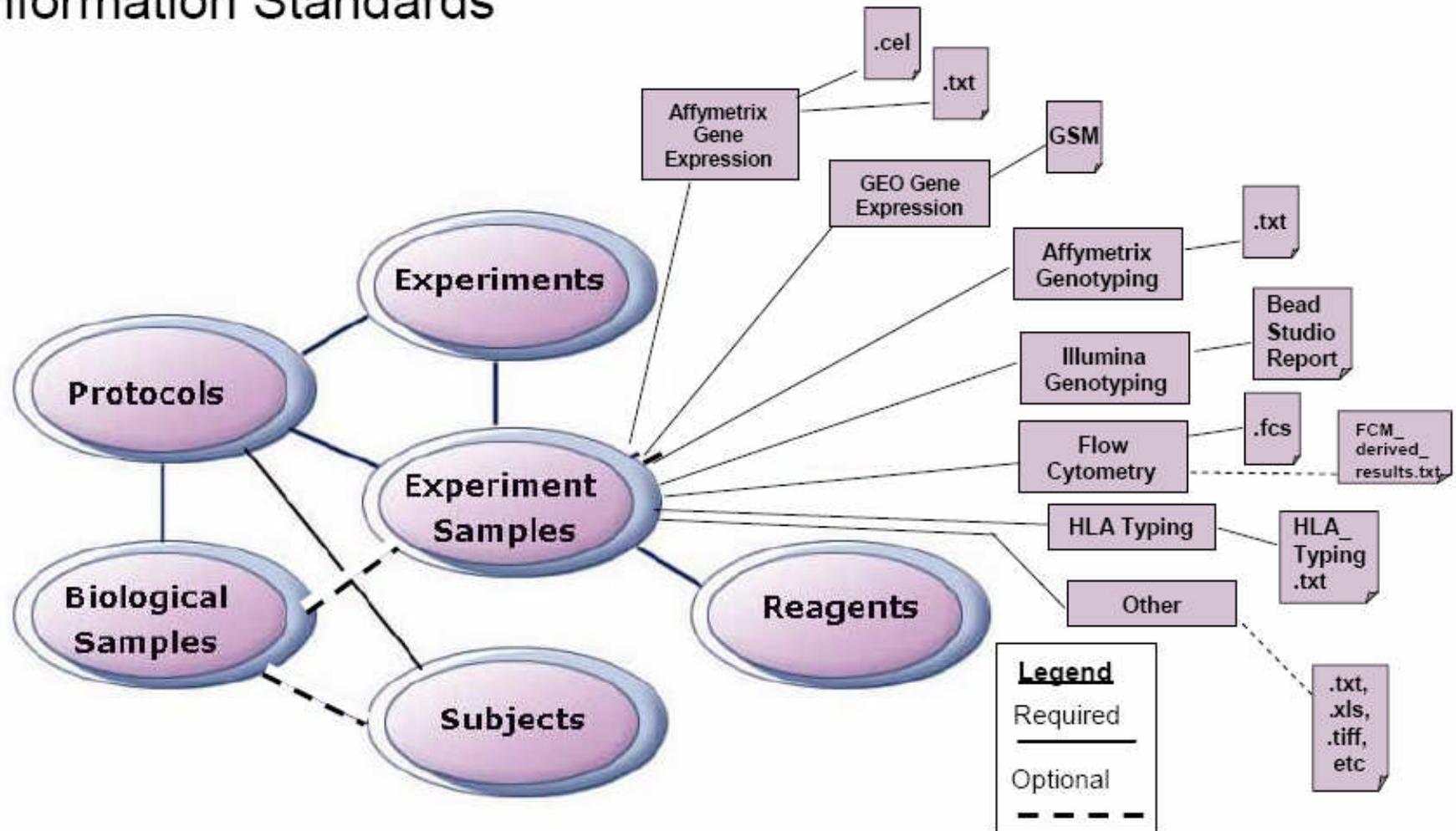
IMMPORT RESEARCH MODEL CLASSES



———— Required Links

- - - - Conditional Links

Results Data File Formats: Post-DAIT Minimum Information Standards





IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

ImmPort Data Submission Process

#2: DOWNLOAD AND FILL THE TEMPLATES

1. Download all the spreadsheet templates you want to use and save them to a local directory.
2. If you are interested in using the XML templates, please contact us.
3. Fill out all the Microsoft Excel file templates and then save them as tab-delimited text files.

***** DO NOT change the original names of the templates.** For instance, **protocols.xls** becomes **protocols.txt** and **NOT protocols-2.txt**.

[Let us look at each of the following templates and see how they relate to each other]



IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

ImmPort Data Submission Process

- The column headers are used to ensure that the data the user enters is properly processed.
Do not edit or delete the column headers.
- **Non-shaded columns are required fields** to be completed and **shaded columns are optional fields.**
- To successfully load data into ImmPort, only the required fields in the templates need to be filled out. However, in order to fully appreciate ImmPort's research data query capabilities, we encourage you to provide as much information as possible for the optional fields.
- Explanatory comments which include the type of information and the size of text to enter into a cell are available by moving the cursor (mouse) over the column header name. Commented column headers have a small red triangle in the upper right corner of the cell.
- Drop lists for some columns (e.g. measurement technique in experiments.xls) are activated by clicking on a cell in the column (i.e. controlled vocabulary).
- The User-Defined ID is intended to be a short name for a metadata instance and must be unique within the scope of a research project.
- When working on the Excel templates, we recommend that you save these templates in Microsoft Excel format until you are ready to upload a submission package. At that point, you will save all templates as tab-delimited text file files.
- The Excel templates "talk" to each other via key linkages (e.g., the Protocol IDs defined in Protocols.xls will be referenced in all the other templates when you indicate which protocol was used for a particular biological sample). These key linkages should match up when used.



IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

ImmPort Data Submission Process

protocols.xls template structure

Required columns:

User-defined ID*: This is the unique identifier that links this template to the other templates like biosample.xls, experimentSample.xls, experiments and subjects.

Protocol File Name*: This is the name of the accompanying file. This can be a word, pdf, excel files etc. **Note:** File names should have NO SPACES.

Optional columns:

Protocol Name: Name that you want to give your protocol in free text.

Protocol Type: Terms selected from a controlled vocabulary such as assay, biomaterial transformation etc.

Protocol Summary: This is a summary of the protocol in free text.

[Let us go ahead and fill out the protocols.xls template]



bioSamples.xls template structure

Required columns:

Biological Sample User-Defined ID*: Unique identifier that links this template to the other templates like biosample.xls, subjects and protocols.

Biological Sample Type*: Terms selected from a controlled vocabulary such as cell, DNA, tissue etc.

Biological Sample Protocol User-Defined ID*: Protocol used in biosample prep. Links to the protocols.xls template.

Optional columns:

Biological Sample Name: Name that you want to give your biosample in free text.

Biological Sample Description: Description of the biosample used.

Biological Sample sub-type: Used to refine the biosample analysis.

Source Biological Sample User-Defined ID: Source of the biological sample e.g. blood.

Source Subject User-Defined ID: Subject the biological sample was derived.

Pooled Sample: Terms selected from controlled vocabulary, Y or N.

Biological Sample Purity: Purity measurement of biological sample.

Biological Sample Concentration: Concentration of the biological sample.

Biological Sample Volume: Volume of the biological sample.

Biological Sample Weight: Weight of the biological sample.

[Let us go ahead and fill out the biosample.xls template]



experiments.xls template structure

Required columns:

Experiment User-Defined ID*: Unique identifier that links to Protocols and is linked from experiment sample.

Experiment Type*: Terms selected from a controlled vocabulary such as cell, DNA, tissue etc.

Measurement Technique*: Terms selected from a controlled vocabulary such as Array, ELISA, clinical trial

Protocol User-Defined ID*: Protocol used in biosample prep. Links to the protocols.xls template.

Optional Columns:

Experiment Name: Name that you want to give your Experiment in free text.

Description: Description of the Experiment.

Hypothesis: Hypothesis set when performing the experiment.

Rationale: The reason for doing the experiment.

Keywords: Keywords of the experiment.

Quality control measures: Any measures taken to make ensure expected results

Experimenters: Who were involved in carrying out the experiment?

Links to publications: What publications are available for this experiment.

Constants: What are some of the variables that were constant in the experiment.

Conditional variables: What variables were conditional for the experiment.

Responding variables: What variables changed from the set experiment constants.

[Let us go ahead and fill out the experiment.xls template]



experimentSample.xls template structure

Required columns:

Experiment Sample User-Defined ID*: Unique identifier that links this template to the other templates like biosample, subjects, protocols, reagents, result files.

Experiment User-Defined*: Experiment to which the experiment sample belongs.

Protocol User-Defined ID*: Protocol used in biosample prep. Links to the protocols.xls template.

Reagent User-Defined ID*: Reagents associated with the experiment submission package.

Optional columns:

Experiment Sample Name: Name that you give your experiment sample in free text.

Experiment Sample Description: Description of the experiment sample used.

Result File or Folder Name: Name of the result file or folder.

Replicate Group ID: Source of the biological sample e.g. blood.

Biological Sample User-Defined ID: Biological sample used in the experiment.

[Let us go ahead and fill out the experimentSample.xls template]



subjectsHuman.xls template structure

Required columns:

Subject User-Defined ID*: Unique identifier that links this template to the other templates like biosample, protocols, reagents, result files.

Species Name*: Species that you are dealing with.

Protocol User-Defined ID*: Protocol used in biosample prep. Links to the protocols.xls template.

Optional columns:

Subject Description: Description of the subject used in the experiment.

Affection Status: Affection level selected from a list of controlled vocabulary.

Affection Phenotype: Translating the affection status to a physical attribute like susceptible.

Enrollment age: The cut off age that the subjects have to be in the experiment.

Gender: Gender obtained from a list of controlled vocabulary.

Ethnicity: What ethnic group does the subject belong?

Race: What race does the subject belong?

Birth place: Where was the subject(s) born?

Population: Which population does the subject belong?

Ethnicity: What ethnic group does the subject belong?

Religion, Language: What is the subjects religion, language?

Family, Subject, Mother, Father pedigrees IDs: What is the pedigree information?

Treatment information: What is the pedigree treatment information?

[Let us go ahead and fill out the subjectHuman.xls template]



Stacked templates and relationships

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	Biological Samples	Version 2.3																	
2	Please do not delete or edit this column																		
3	Column Name	Biological Sample User-Defined ID*	Biological Sample Type*	Biological Sample Protocol User-Defined ID*	Biological Sample Protocol ImmPort Accession*	Biological Sample Name	Biological Sample Description	Biological Sample sub-type	Source Biological Sample User-Defined ID	Source Biological Sample ImmPort Accession	Source Subject User-Defined ID	Source Subject ImmPort Accession	Pooled Sample	Biological Sample Purity	Biological Sample Concentration	Biological Sample Volume	Biological Sample Weight	treatment name	
4		BSample_1	Tissue	Golden Gate Assay		PatientSar	Blood extract	Blood											
5		Sample_1	DNA	Smallpox Vaccination		PatientSar	DNA extract	DNA	BSample_1										

	A	B	C	D	E	F
1	Protocols	Version 2.3				
2	Please do not delete or edit this column					
3	Column Name	Protocol User-Defined ID*	Protocol File Name*	Protocol Name	Protocol Type	Protocol summary
4		Golden Gate Assay	GOLDENGATE_ASSAY_FINAL.pdf	Golden Gate Assay1	Assay	This is summary of a gate assay
5		Smallpox Vaccination Method	smallpox_vaccination_method.pdf	Smallpox Vaccination Method1	Assay	This is summary of a vaccination method

Note: The color codes are used in the templates to guide in linking metadata instances together

Once all the templates have been filed out, convert all the files as tab-delimited files and then zip them with all the accompanying and relevant files. In order to validate all the files, use the Data Package Validator which has been described elsewhere.



IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

ImmPort Data Submission Process

DAIT Minimum Information Standards ensures that:

- i. Links to Protocols from Subjects, Biological Samples, and Experiment are made.
- ii. Samples, and Experiments are required.
- iii. Links to Reagents from Experiment Samples are required.
- iv. Reagent descriptions will require inclusion of manufacturer and catalog number.
- v. For ELISA, ELISPOT, MBAA results file, it is optional to send the file in but it is recommended to submit it with Experiment Sample. Submit file only once.

Referencing Protocols

Please describe to ImmPort your protocols. Attach a file of any format (.doc, .xls, .pdf, etc).
Add a reference to your protocol. Complete one additional column in a template.
A single protocol can be referenced that covers all aspects of the experiments.

However, selected assay results will have enhanced inclusion criteria:

Affymetrix Gene Expression results should include .cel and .txt files.
Flow Cytometry results should include a .fcs file.

Referencing Reagents

Many commercially available gene expression and genotyping arrays are pre-loaded into ImmPort and can be referenced as reagents.



IMMPORT
BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY

ImmPort Data Submission Process

The Consequences for Data Submission:

- The column headers are used to ensure that the data the user enters is properly processed.
Do not edit or delete the column headers.
- **Non-shaded columns are required fields to be completed but the shaded columns are optional fields.**
- To successfully load data into ImmPort, only the required fields in the templates need to be filled out. However, **in order to fully appreciate ImmPort's research data query capabilities, we encourage you to provide as much information as possible for the optional fields.**
- Explanatory comments which include the type of information and the size of text to enter into a cell are available by moving the cursor (mouse) over the column header name. Commented column headers have a small red triangle in the upper right corner of the cell.
- Drop lists for some columns (e.g. measurement technique in experiments.xls) are activated by clicking on a cell in the column (i.e. controlled vocabulary).
- The User-Defined ID is intended to be a short name for a metadata instance and must be unique within the scope of a research project.
- When working on the Excel templates, we recommend that you save these templates in Microsoft Excel format until you are ready to upload a submission package. At that point, you will save all templates as tab-delimited text file files.
- The Excel templates "talk" to each other via key linkages (for example, the Protocol IDs defined in Protocols.xls will be referenced in bioSamples.xls when you indicate which protocol was used for a particular biological sample).
It is important to make sure all of these key linkages match up when they are used.



#3: SAVE TEMPLATES, VALIDATE AND UPLOAD THE PACKAGE

After you have successfully filled out your templates and results files correctly,

3. Save each completed template as tab-delimited text file and collect/prepare all accompanying files in a single folder. Zip them using WinZip. Make sure that you select all the file contents and Zip them. **Do not zip the folder.**
4. In order to ensure that all the completed templates are in the correct format and content, use the Data Submission Package Validator by going to Step 2. Check Data in .Zip file. A movie on how to use the data submission package Validator has been given elsewhere.
5. Correct any issues in your package. If there are no more errors, submit your data into ImmPort.
6. Under Submit Data, go to Step 3. Send Data in .Zip file.
 - i. If a package is less than or equal to 1 Gigabyte – You can submit online
Check upload online check box
Select the file using the browse button
Click submit
Record the registration number you receive
 - ii. Alternatively, if file is greater than 1 Gigabyte – You CANNOT submit online as in # I
Uncheck upload online check box
Click submit
Record the registration number (Upload Ticket #) you receive
Contact the helpdesk (with registration number) to arrange for file delivery
7. Click Step 4. Review Submission Status and Results to track your package in ImmPort. Select your project and click '**Set Filter**'. A '**Completed**' status indicates the package successfully uploaded without errors. An email notification is send the submitter with a hyperlink to the Data Submission History page.



Data Submission / Data Submission Details

[Submit Data](#) | [Submission History](#) | [Resources](#) ▾

[Back to History](#)

[Data Management Home](#)

Data Submission Details

Ticket Number: angelst_20090817_90
 Submitted Zip File: EGFP-ELISA_pkg.zip
 Grant/Contract Number: HHSN2662004000765
 Project Title: TESTING: Bioinformatics Integration Support Project
 Data Format: ELISA-MBAA_report_file
 Status: Completed ([download report](#))
 Submitter: angelst
 Submit Date: 08/17/2009 08:22:23
 Metadata File Format: TAB_DELIM
 Notes: EGFP ELISA data from DNA vaccine testing.

Data Parsing/Loading Results

12 items found, displaying all items.

Result ID	File Name	Status	File Parse Exception	Description
2394	EGFP-ELISA_pkg.zip	Pending		server name:testapp2.import.net
2395	EGFP-ELISA_pkg.zip	Started		
2396	VP22DNAvaccineStudy.doc	Archived		size: 43008 bytes
2397	mouse-serum-ELISA.doc	Archived		size: 35328 bytes
2398	protocols.txt	Parsed		
2399	reagents.txt	Parsed		
2400	subjectsAnimal.txt	Parsed		
2401	bioSamples.txt	Parsed		
2402	experiments.txt	Parsed		
2403	experimentSamples.txt	Parsed		
2404	EGFP-ELISA_titers.xls	Archived		size: 13824 bytes
2405	angelst_20090817_90__EGFP-ELISA_pkg.zip	Completed		

12 items found, displaying all items.