pISA-tree: Standard project directory tree

A. Blejec, ... , K. Gruden

2016-11-21

# What is pISA-tree?

This set of batch files is used to create standard directory tree for research projects.

## Layers

At the top layer, you have to create a directory (pISA\_projects root directory) which will contain future projects. The root directory is the place for initial batch files. Step one is to copy/download the batch files from the [[set source]] into the root directory.

**Project** is organized as a collection of one or more **investigations**. An **investigation** is similarly organized as a collection of one or more **studies**. Each **study** has it's own collection of one or more **assays**. Assays can be of specific type (e.g. conected to MicroArray, NGS, Modeling, Statistical analysis, ...) and are structured accordingly.

## Batch files

* makeproject.bat - makes new **project** directory tree
* makeInvestigation.bat - makes new **investigation** directory tree (subdirectory tree within the **project**)
* makeStudy.bat - makes new **study** (subdirectory tree within the **investigation**)
* makeAssay.bat - makes new **assay** (subdirectory tree within the **study**)

# Creation of the directory tree

The directory tree is a way to enforce the subordination of pISA layers. To emphasize the layer type, directory names are constructed automatically using the standard prefix and short layer ID. Standard prefixes are:

* \_p\_for project
* \_I\_for investigation
* \_S\_for study
* \_A\_for assay

## project

To create a new project, run (double click) the file makeProject.bat and enter the project short name (ID). This will make a direcory tree, description files and a local copy of makeInvestigation.bat. Short project name (ID), automatically prefixed with \_p\_, is used as the name of the directory (no spaces or special characters, file name conventions apply!!). If you set the project ID as bla the project directory name will automatically become \_p\_bla.

## Investigation

To create a new investigation, run (double click) the file makeInvestigation.bat and enter the study short name (ID). This will make a direcory tree, description files and a local copy of makeStudy.bat. Short investigation name (ID), automatically prefixed with \_I\_, is used as the name of the directory (no spaces or special characters, file name conventions apply!!). The investigation directory name for the investigation ble will be \_I\_ble.

## Study

To create a new study, run the makeStudy.bat and enter the study ID (Short name). This will make a direcory tree with several standard folders, description and auxiliary files and a local copy of makeAssay.bat. The study folder name will be \_S\_blu for a study with short name (ID) blu.

## Assay

Analyses for each study are collected in the folder of that study. To make a new assay, run the makeAssay.bat file.

You will be asked to enter assay Class [Wet/Dry]:

* Wet: measurements on the biological material (MicroArray, NGS, PCR, ...)
* Dry: process data (Statistcs, Modelling, ... )

Then you will enter the assay Type (see documentation for covered types: NGS, MA, STat ...) and assay ID (Short name, for example ASSAY1). Short assay name and type (separated by '.'), prefixed by \_A\_ are used as the name of the assay directory tree (for expample: \_A\_ASSAY1.NGS).

Folders in assay directory trees for different Classes slightly differ, according to the need of the specific Class.

## Description files

Each level has a *description* file, listing the informative items for that level. Description files are tab delimited files with two columns:

1. item name (ended by a colon)
2. item description or value

Item description can be some text (for example investigator's name or a longer description of the study, study, and analysis) or a value (for example the path to *phenodata* file). Each item should be typed in one line. Special escaped characters for line break (\n) and tab (\t) are allowed. Be careful if the description contains prime symbol (' ,as in 5'), it is safer to spell it, like 5-prime.

An example of the description entry is:

Investigator: Miha Mihav  
Phenodata: ./data/phenodata.txt

The project level directory contains the file common.ini, which is organized in the same way as description files. This file contains fields and descriptions that do not change, for example principal investigator name. Information in this file will be appended to the description files for investigations, studies, and assays.

# Auxiliary batch files

## showDescription.bat

Collects all description files in a tree below the current level. Descriptions are typeset in the file **DESCRIPTION.MD**. (plain text file in a markdown format)

## xcheckDescription.bat

Checks all description files for missing required information (*)in a tree below the current level. Produces file* ***xCheckDescription.md*** *which is similar to one produced by showDescription.bat but lists only lines with asterisks (*).

## showTree.bat

List a directory tree below the current level in the file **TREE.TXT**.

## update.bat

Replaces batch files in existing tree with the updated versions in the project directory. Use after update from [GitHub](https://github.com/ablejec/pISA/)