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 (project root directory) which will contain future investigations. 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.

**Investigation** is organized as a collection of one or more **studies**.  
Each **study** has it's own collection of one or more **assays**. Assay can be of different type (e.g. conected to MicroArray, NGS, Modeling, Statistical analysis, ...) and are structured accordingly.

## Batch files

* makeInvestigation.bat - makes new **investigation** directory tree
* 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

## 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 study name (ID) is used as the name of the directory (no spaces or special characters, file name conventions apply!!). Each study tree has a \_STUDIES subdirectory, where investigation studies will be rooted.

## Study

Investigation studies are collected in the subdirectory \_STUDIES. In the main \_STUDIES directory, you will find the makeStudy.bat. To create a new study, run the makeStudy.bat and enter the study ID (Short name). This file will make a new directory tree, rooted in the \_STUDIES subdirectory. Each study tree has a \_ASSAYS subdirectory and a copy of the makeAssay.bat file.

## Assay

Analyses for each study are collected the \_ASSAYS 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 '-') are used as the name of the assay directory tree (for expample: ASSAY1-NGS).

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

For example:

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

The study 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/)