Overview of data (types) in the Matlab model

molecules (species in SBML)

molecule	**attributes**
Transcriptional Units	ro | rw
RNAs	ro | rw
Protein Monomers	ro | rw
Protein Complexes	ro | rw
Metabolites	ro | rw | req | roreq

ro read only
rw read/write
req requirement
roreq read only + requirement

parameters

- fixed for the simulation
- see also
  https://simtk.org/project/xml/downloads.xml?group_id=714
  https://simtk.org/frs/download.php?file_id=3757

variables (state)

- 83 states
  states *MetabolicReaction*, *Metabolite*, *ProteinComplex*, *ProteinMonomer*, *Rna*, *Stimulus* are separately described in more detail
- changed (written) during simulation by at least one process
- see also
  https://simtk.org/project/xml/downloads.xml?group_id=714
  https://simtk.org/frs/download.php?file_id=3759
  https://simtk.org/frs/download.php?file_id=3760

Naming scheme

molecules (species in SBML)

- connect the MoleculeName by double underscores ('__') with a CompartmentName and/or a MoleculeState
  MoleculeName__CompartmentName__MoleculeState
  MoleculeName__CompartmentName
  MoleculeName__MoleculeState
- list of CompartmentNames
  - m (membrane)
  - e (extracellular)

  The cytosol ('c') can be ignored. If no CompartmentName is used it is assumed the molecule is in the cytosol.

- list of possible MoleculeStates
  - aminoacylated
  - bound
  - damaged
  - folded
  - inactivated
  - intergenic
  - mature
  - misfolded
  - nascent
  - processed
  - processedI
  - processedII
  - signalsequence

  from StatePropertyRowColIDs.xlsx (https://simtk.org/frs/download.php?file_id=3760)

  The state 'activated' can be ignored. If no MoleculeState is used it is assumed the molecule is in the state 'activated'.

  If you have to use other states please don't use a dash ('-') in the state name. This might cause problems in the further processing of the files.

- examples: ATP, ATP__e, MG_001_MONOMER__inactivated, MG_001_MONOMER__m__inactivated

- MoleculeNames can be found in five lists (files) on GitHub https://github.com/whole-cell-tutors/wholecell/tree/master/wholecell-integration
  - Molecules_names_TUs.csv

    - Molecules_names_RNAs.csv
    - Molecules_names_Monomers.csv
    - Molecules_names_Complex.csv
    - Molecules_names_Metabolites.csv

- make a local copy of these files for each process and update the files
  example: Molecules_names_Metabolites_translation.csv

  put an ‘x’ in the respective column
  you can remove lines from these files and just keep the ones for your process

parameters

If a model needs a parameter the team who is implementing this model has to take care of this parameter (the Integration team won’t provide it, sorry!).

variables (state)

- connect the StateID by double underscores (‘__’) with the PropertyName
  StateID__PropertyName

- example: StateID *Metabolite* with PropertyName *counts* -> variable *Metabolite__counts*

- VariableNames can be found in a list (file) on GitHub
  https://github.com/whole-cell-tutors/wholecell/tree/master/wholecell-integration
    - Variables.csv

- make a local copy of this file for each process and update the file
  example: Variables_translation.csv

  you can remove lines from this file and just keep the ones the process changes (writes) or you put an ‘x’ in the fourth column (*Changed*)

	A	B	C	D
1	**Variable**	**State ID**	**Property name**	**Changed**
2	Chromosome__abasicSites	Chromosome	abasicSites	
3	Chromosome__complexBoundSites	Chromosome	complexBoundSites	
4	Chromosome__damagedBases	Chromosome	damagedBases	
5	Chromosome__damagedSugarPhosphates	Chromosome	damagedSugarPhosphates	
6	Chromosome__doubleStrandedRegions	Chromosome	doubleStrandedRegions	

requirement parameters

- make sure your model has a requirement parameter for each molecule (species) that is a requirement

- the requirement itself has to be calculated by the model (the Integration team won't do it for you, sorry!)
- connect the name you have made up for a molecule by two underscores ('__') with the text *requirement*
- example: ATP__requirement or ATP__e__requirement