OVERVIEW

In order to perform section registration, NeuroInfo requires atlas data to be presented in a specific form. All of the most important information is stored in an XML file, which defines the axes and their orientations, as well as the file paths where supporting files are stored.

One of these supporting files (a CSV file) contains the names of and relationships between the anatomical regions described by the atlas. Each region is assigned a unique integer value as an ID. Regions are described in a hierarchy; if a particular region is contained completely within a larger region, then the larger region is said to be the "parent" of the smaller region. At the top of the hierarchy is always a "region" which contains the entire brain – this will be the parent of all of the largest regions which are not contained within any other regions.

The atlas also requires a set of three images which are used in different ways. Firstly, there is a reference image – this is not used for any detection or quantification. Rather, it is the image displayed to the user for the purposes of navigation and visualization. The other two images are used for detection and registration. The anatomy image is a "paint by numbers" version of the reference image, where each pixel's value matches the unique ID value of the region in which it is located. The mask image divides the image into "tissue" and "not-tissue" by coloring all pixels that are outside the tissue with black and coloring all pixels that are inside the tissue with white.

Finally, a set of VTK files containing 3D models of all of the regions described in the atlas is required. This data is used to derive the shape of the cross-sectional area of each region as they appear within registered slices. Each region described by the ontology CSV must have a corresponding VTK file that contains the 3D model of that region.

In summary, a NeuroInfo atlas is comprised of:
1. An XML file defining the coordinate system and including references to supporting files
2. A CSV file containing ontological information
3. A set of reference image stacks in NRRD format
4. A set of VTK files containing 3D models of the regions/subregions described by the ontology

DETAILS: ATLAS STRUCTURE

1. <u>Coordinate System and References to Supporting Files (.xml):</u>
   - \<atlas> - the root element for all atlas data
     - o \<name> - a custom name of the atlas
     - o \<version> - the version of the atlas
     - o \<description> - a custom description of the atlas
     - o \<orientation> - for each axis, defines the direction along the axis in which the values increase toward positive numbers, using conventional medical imaging terminology
       - ▪ \<x-axis> - choose "left" or "right", where "left" is the subject's left and "right" is the subject's right

- - **\<y-axis\>** - choose "dorsal" or "ventral", where these refer to the dorsal and ventral surfaces of the brain, respectively
    - **\<z-axis\>** - choose "rostral" or "caudal", where these refer to the rostral and caudal directions within the brain, respectively
  - **\<reference-image\>** - there are actually two images used for reference: one showing the histology to be displayed to the user, and another as a binary mask delineating where the tissue is and is not
    - **\<name\>** - a custom name for the image used (does not need to match the actual filename of the reference image)
    - **\<modality\>** - a custom description of the modality used to acquire the image (ex. "two-photon fluorescence")
    - **\<image\>** - the relative path to the histology reference image file (this is the image that the user will see)
    - **\<mask\>** - the relative path to the masked version of the reference image file (this will not be displayed to the user)
  - **\<annotation\>** - information about the delineation of regions and supporting files
    - **\<anatomy-image\>** - the relative path to the image where each pixel's value corresponds to the ID of the region/subregion in which that pixel lies (ontology .csv, column 0/A)
    - **\<mesh-path\>** - the relative path to the directory which contains the 3D model files (.vtk) of all regions/subregions
    - **\<region-ids\>** - a comma-separated list containing all unique IDs of all regions/subregions (ontology .csv, column 0/A) for which there is at least one pixel in the anatomy image that uses the ID to delineate a region/subregion, in increasing order. This property is optional, but including it will allow some operations to progress faster.
  - **\<ontology\>** - the relative path to the ontology .csv
  - **\<midsagittal-plane\>** - specifies the location and orientation of the midsagittal plane within the atlas. This property is optional and may be omitted (e.g. in cases where the atlas is asymmetrical and cannot be cleanly divided by a single plane)
    - **\<point\>** - a point which lies on the midsagittal plane
    - **\<normal\>** - a unit vector normal to the midsagittal plane and pointing along the left direction of the X axis.

2. <u>**Ontology (.csv):**</u>
   - The first cell of each column (i.e. in row 0) is expected to be the title of the column and will not be used
   - Each subregion must be defined below its parent region/subregion (i.e. the row number at which the subregion is found must be greater than the row number at which its parent is found)
   - AT MINIMUM: there must be EXACTLY ONE root region which contains the entire brain, including all regions/subregions. It MUST have:
     - Name "root" (in column 2/C)
     - Short Name / Acronym "root" (in column 3/D)
     - NO Parent ID (in column 8/I).
   - The data for each region/subregion must be organized in the following manner:
     - For column 0 (or "A"): ID

- Each region/subregion (including the "root") must have unique positive integer ID
  - o For column 2 (or "C"): Name
    - A full name that clearly identifies the region/subregion, without any abbreviations (i.e. as it would appear when it is first used in a publication)
  - o For column 3 (or "D"): Short Name / Acronym
    - An abbreviated name that also refers to the region/subregion (ex. for "Subventricular Zone", this might be "SVZ")
  - o For column 8 (or "I"): Parent ID
    - Each region/subregion must have EXACTLY ONE parent: the region/subregion that contains this subregion. The parent's ID (from the parent's column 0/A) should be given here. If this region is at the top of the ontology hierarchy and is one of the first, large divisions of the brain with no anatomical region that contains it, then its parent should be given as the ID of the "root" region. The "root" region itself MUST have NO Parent ID listed here.
  - o For column 13 (or "N"): Hexadecimal Color
    - Each region/subregion (including the "root") must have a color specified as a hexadecimal triplet. These do not necessarily need to be unique, but we advise that, if colors are to be repeated, then sibling subregions (ie. subregions which share the same parent) should share the same color.
  - o For all other columns: User-Defined
    - Feel free to use any of the other columns not listed here for your own purposes – they will not be imported into the atlas.

3. <u>Reference Images (.nrrd):</u>
   1. Histology reference image: An image stack showing the histology which the user will view in order to calibrate/register other images to the atlas
   2. Histology image mask: An image stack with a binary mask that indicates where there is tissue (full intensity) and where there is empty space (zero intensity) in the histology reference image
   3. Anatomy reference image: An image stack that uses the colors specified in the ontology .csv (column 13/N) as hexadecimal triplets to delineate each region/subregion visible in the histology reference image
   - Images must be unsigned image stacks
   - Images must NOT have any interpolated values
   - Images may have a maximum bit depth of 32 bits

4. <u>3D Models (.vtk):</u>
   - Each region/subregion listed in the ontology file must have a corresponding .vtk file
   - The .vtk file of a region/subregion must be named according to the Short Name / Acronym (ontology .csv, column 3/D) and ID (ontology .csv, column 0/A) of the region/subregion:
     - o The format of the .vtk filename is: "[Short Name / Acronym]_[ID].vtk"
       - IMPORTANT: if the Short Name / Acronym uses the forward slash "/", then each forward slash should be replaced with an underscore "_" when naming the .vtk file

- The model must use the Polydata type, and must be defined using Points and Triangle Strips

EXAMPLE: ADDING THE WAXHOLM ATLAS

1. <u>XML Reference:</u>
   - This file should be placed in C:\ProgramData\MBF Bioscience\Atlases, and all of the other files mentioned below may be placed in the same folder, or may be placed in any subdirectory of this folder
   - Make sure to edit the paths in this file so that they contain the correct relative paths to the other files
     - Ex: If the Ontology file is called "MyAtlasOntology.csv" and is placed in the Atlases folder (i.e. in the same place as the XML Reference file), then <ontology> in the XML Reference file must be changed to "MyAtlasOntology.csv". If instead a folder named "MyAtlasFiles" is created inside the Atlases folder and the Ontology file is placed inside that folder, then <ontology> must be changed to "MyAtlasFiles\MyAtlasOntology.csv"
   - Orientation:
     - X axis: right
     - Y axis: rostral
     - Z axis: dorsal
   - Region IDs: this atlas doesn't contain as many regions as the Allen Institute atlas, so while it's okay to include the list of region IDs, it likely won't speed up performance much
   - Midsagittal Plane: this atlas is not symmetric and therefore does not require a midsagittal plane to be defined

2. <u>Ontology:</u>
   - The label file WHS_SD_rat_atlas_v2.label contains IDs, names, and colors of regions (this file can be viewed in Notepad)
   - The first column contains unique IDs for each region
     - These can be directly copied into column 0/A of the ontology .csv
   - The second, third, and fourth columns are the red, green, and blue values, respectively, of the color used for each region
     - These three separate values must be converted to a hexadecimal triplet before being inserted into column 12/N of the ontology .csv (ex. 255, 52, and 39 convert to FF3427)
   - The final column contains the name of each region
     - These can be directly copied into column 2/C of the ontology .csv
   - A "root" region should be added with:
     - A unique ID that hasn't been used by any other regions yet
     - The name and short name / acronym "root"
     - No parent
     - The color "FFFFFF"
   - Since the data is not defined as a hierarchy and instead only contains the smallest divisions, a neuroanatomist will need to go through the data and determine:
     - Which short name / acronym to use for each region

- Which parent regions need to be added – this requires adding a new row to the ontology .csv and choosing values for each of the columns
- Which parent region each of the regions belongs to (remember: the largest divisions should have the ID of the "root" region as the parent ID, and the "root" region should have no parent ID at all)

3. <u>Images:</u>
   o Anatomy image: WHS_SD_rat_atlas_v2.nii.gz
   o Reference image: WHS_SD_rat_T2star_v1.01.nii.gz
   o Mask image: the anatomy image must be passed through a thresholding operation, where all pixels with a value of 0 remain 0, and all other pixels will get a value of 255
   o The images provided by Waxholm are in NIfTI format and will need to be converted to NRRD format before using with NeuroInfo (this can be done using FIJI-ImageJ)

4. <u>3D Models:</u>
   o This atlas doesn't contain 3D models – they will need to be built using the data already available in this atlas (ex. take the anatomy image and ontology .csv together, and use that information to determine the shape and size of each 3D model that needs to be created)