Appendix 1: Overview of Iolite's internal structure and data format

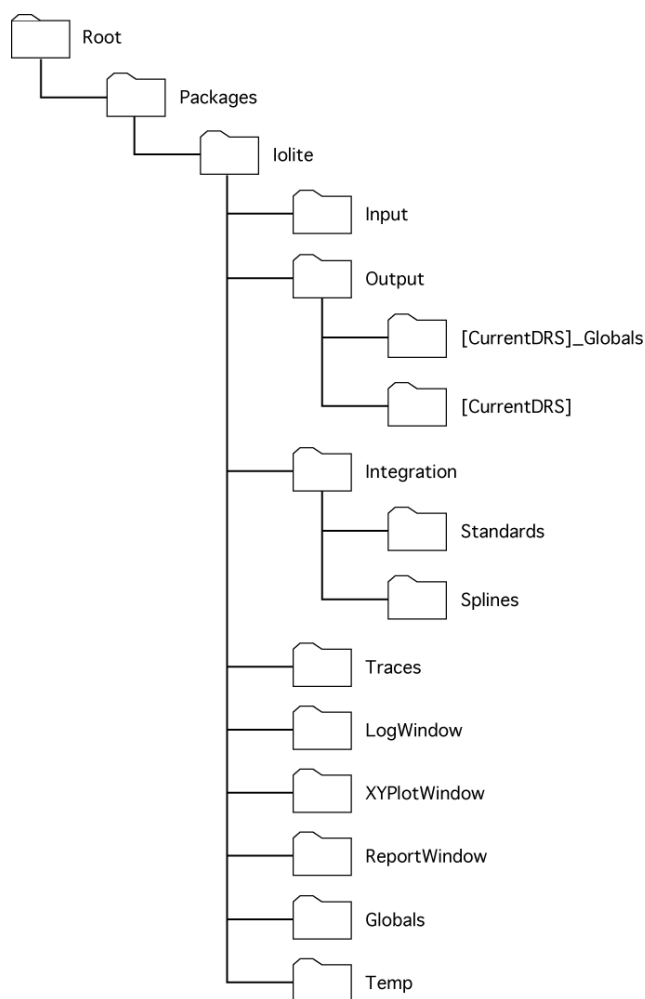**1   Iolite's internal data structure**

Figure A1 - Schematic of Iolite's internal data folder structure

The internal data structure of Iolite has been organized using IgorPro's internal data folder system – these folders are similar to conventional file folders on a hard disk, but they are only stored within IgorPro, and do not exist outside of the experiment file in which they were created. The data folder system has been used in Iolite to subdivide the numerous waves (waves are data arrays of up to 4 dimensions, e.g., input channels are an example of a 1-D wave), variables and strings created during normal data processing into numerous subfolders, thereby allowing easier navigation of the underlying data. Figure A1 contains a schematic illustration of Iolite's data folder hierarchy, and the purpose and contents of each folder are summarized below:

Input – Contains all input data waves created during the import of machine data.

Output – Contains subfolders for each DRS that has been active in the current experiment. This folder also contains the "ListOfIntermediateChannels" and "ListOfOutputChannels" global strings that are used to store information about which waves have been created by the DRS.

[CurrentDRS] – This folder will have a name identical to the relevant DRS, if more than one DRS has been used in the current experiment then each will have a corresponding folder. This allows the user to switch between DRS within an experiment without any loss of information or settings specific to the DRS. All waves created by the DRS during data processing are stored in this folder, as well as any other waves, global strings or global variables required by the DRS (e.g. the interactive down-hole curve fitting windows of the "U_Pb_Geochronology" DRS).

[CurrentDRS]_Globals – Each CurrentDRS data folder has this associated folder, which contains global strings and variables to store all settings for the DRS contained in the header of the DRS procedure file, and modified by the user via the "Edit Settings" button on the Main Control Window.

Integration – The 3-dimensional matrices used to store information for integration types are stored in this data folder. There are two matrices for each integration type: a numeric and a text matrix (more details are provided in Section 2 of this appendix). In addition, it contains relevant waves, global strings and global variables required for maintaining the Main Control Window.

Standards – Stores information for each of the reference standard materials in separate 2-dimensional waves. The data are read from disk when opening a new experiment in IgorPro, or when "Update Old Experiment" is run.

Splines – This data folder is used solely to store the splines generated automatically by Iolite for each integration type, in combination with each of the input, intermediate and output waves.

Traces – Contains all relevant waves, global strings and global variables for maintaining the Traces Window interface.

LogWindow – Contains all relevant waves, global strings and global variables for maintaining the Log Window interface.

XYPlotWindow – Contains all relevant waves, global strings and global variables for maintaining the XY Plot Window interface.

ReportWindow – Contains all relevant waves, global strings and global variables for maintaining the Report Window interface.

Globals – All Iolite settings independent of the DRS are stored in this folder, and can be modified by the user via the "Edit Settings" button on the Main Control Window.

Temp – Used to store waves, global strings and global variables that are temporarily required for a specific function (e.g. an interactive window), but do not need to be stored for later use. The Temp folder is emptied whenever an experiment is saved, thus freeing memory and improving Iolite's performance.

2   Wave (array) formats in Iolite

There are a variety of wave array formats available in IgorPro, with the following characteristics:

Single precision – This is the default wave type in IgorPro, it stores information as single precision floating point (4 byte, 32 bit) numbers, which corresponds to approximately 7 significant figures of precision. Although this format is usually sufficient for intensity and ratio calculations in Iolite (i.e. channels), it is not sufficient to store date/time information with sub-second precision, and should never to used for time waves.

Double precision – These waves have double the precision of single precision waves (8 bytes, 64 bits), which corresponds to approximately 15 significant figures of precision. Although this format consumes twice the memory of single precision waves, it is necessary for the accurate representation of absolute times in Iolite, and is therefore used for all time waves.

Text – Text waves can be used to store strings of text instead of numbers. There are no restrictions on the length of a string, or the characters it contains, meaning that a single point in a text wave can be further divided into separate strings using a list (i.e. a series of strings separated by a specified character, typically a semicolon).

Multiple dimensions – Regardless of their format, waves in IgorPro can have between one and four dimensions. Although the majority of waves created by Iolite are single dimensional (e.g. time-series data), both two and three-dimensional matrices are also used. Dimensions are referred to in IgorPro as rows, columns, layers and chunks, respectively.

Although it is not necessary to discuss all waves created in Iolite here, the characteristics of the following general types are worth noting:

Input waves (Input Channels) – These are single precision one-dimensional waves, with each having a corresponding time double-precision time wave. Each point of the wave contains the value of the input data (typically measured intensity expressed as counts per second or in volts) for the absolute time (expressed as seconds since January 1, 1904) of the corresponding point in the associated input time wave. The import of data from multiple files is accommodated by concatenating (splicing) new data into existing input waves. The newly imported data can fall before, after, or within (providing there is no time overlap with existing data) previously imported data in the input wave, and need not have the same time spacing as preceding files. However, the data being imported must be temporally ascending, and cannot have two data points with identical times. There is no minimum number of points for input waves, and although users can expect noticeable calculation times if the number of points exceeds 100,000, there is no explicit maximum (tests containing approximately 1 million points have not caused Iolite to crash). Data acquired simultaneously using more than one machine can be imported despite overlapping in time, provided that none of the channels have identical names. Note that data points from a single imported analysis are considered to be coherent, whereas time periods between imported analyses are broken by a single NaN (not a number) data point in order to avoid inappropriate interpolation of values across time gaps.

The Index Time wave – Although input channels are free to each have different absolute time indices, the majority of calculations conducted by a DRS need to be able to compare channels at the same moment in time. Therefore, DRS procedures typically begin by creating an "Index Time" wave using an input wave specified by the "Index Channel" global string in the DRS header. The index time wave is a direct (double precision) copy of the Index Channel's time wave, and will include any gaps in acquisition, it is therefore important to select an Index Channel that has been acquired for all analyses. If data have been acquired simultaneously from multiple machines it is also recommended to select an Index Channel from the machine with the highest sampling frequency to minimize down-sampling of input data.

DRS waves (Intermediate and Output Channels) – These waves are created during execution of the DRS procedure file, and are mapped against the Index Time wave. If an input time wave does not have points directly corresponding to the Index Time wave then adjacent points will be used to linearly interpolate values. Users are free to create as many DRS waves as desired for data processing, and can choose to make any or all of the created waves appear in Iolite's channel selection lists by adding them to the global "ListOfIntermediateChannels" or "ListOfOutputChannels" strings.

Spline waves – These are used for calculating interpolated values for reference standards and baselines, with separate one-dimensional waves used for each channel to be interpolated, for each integration type containing data. A single time wave is created for each integration type (with a number of points determined by the interpolation method used), with the splines of each respective channel mapped onto this time wave. To allow for easier use during data reduction, an Iolite function is provided for interpolating splines onto the Index Time wave. Note that cubic splines and smoothed cubic splines are not true curves, but a series of linear interpolations between individual data points. The spacing of these points is typically 30 seconds, although the minimum and maximum numbers of points in a spline wave are set to 200 and 40000, respectively, meaning that for experiments containing under 100 minutes, or over ca. 2 weeks of data, data point spacing will vary.

Integration matrices – These are three-dimensional waves used to store selected integration periods, and associated statistics calculated from each channel wave using the data points encompassed by each time period. Each integration type used in an experiment has its own numeric integration matrix, as well as a matching text matrix that is used to store non-numeric information (e.g. the selected interpolation method). In both matrices each row stores information for a single integration period, each column stores information for a specific channel, and each layer is used to store a separate type of value. In the numeric matrix the first six layers are specified as follows:
1 Stores the median time in seconds since 1 January, 1904.
2 Stores half of the duration of each integration period, combined with layer 1 this allows convenient graphing of integration periods using error bars.
3 Average intensity.
4 2 standard errors (or equivalent).
5 Number of points after removal of outliers.
6 Number of points rejected as outliers.

In the text matrix, the following layers are also pre-defined:

1 Stores the spline type selected for the integration type.
2 Stores any annotations made for an integration period.

In addition to the above layers, it is possible to append layers to either the numeric or text matrices if required. For example, the uncertainty propagation methods provided in Iolite store calculated values in a new layer of the numeric matrix, instead of overwriting original uncertainties. Note that in order to increase the flexibility of the numeric and text matrices, data are written into, and read from, the arrays using "labels" instead of dimension number.

Image matrices – These are two-dimensional waves used to store the images created in Iolite. Each channel for which an image is created has its own matrix wave, with each point in the array storing the value of that channel for a given x (rows) and y (columns) position.