

# WebPlotDigitizer User Manual

Version 4.6

Ankit Rohatgi\*

November 22, 2022

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	History . . . . .	2
1.2	User Manual and Tutorials . . . . .	2
1.3	License . . . . .	4
1.4	Source Code . . . . .	4
1.5	Availability . . . . .	4
1.6	Supported Browsers . . . . .	4
1.7	Citing WebPlotDigitizer . . . . .	4
1.8	Reporting Issues . . . . .	5
1.9	Data Privacy . . . . .	5
1.10	Funding . . . . .	5
<b>2</b>	<b>Loading Plots</b>	<b>5</b>
2.1	Supported Image Formats . . . . .	5
<b>3</b>	<b>Calibrate Axes</b>	<b>6</b>
3.1	2D (X-Y) Plot . . . . .	6
3.1.1	Calibration Value Format . . . . .	7
3.2	Bar Charts . . . . .	8
3.2.1	Bar Charts vs. Histograms . . . . .	8
3.2.2	Axes Calibration for Bar Charts . . . . .	8
3.3	Polar Diagram . . . . .	9
3.4	Ternary Diagram . . . . .	10
3.5	Maps/Microscope Images . . . . .	11
3.6	Image (Align to Image Pixels) . . . . .	11

---

\*E-Mail: [ankitrohatgi@hotmail.com](mailto:ankitrohatgi@hotmail.com)

<b>4</b>	<b>Grid Removal</b>	<b>12</b>
<b>5</b>	<b>Acquire Data</b>	<b>13</b>
5.1	Manual Mode . . . . .	13
5.2	Automatic Mode . . . . .	15
5.3	Digitization Algorithms . . . . .	17
5.3.1	Averaging Window . . . . .	17
5.3.2	X Step with Interpolation . . . . .	17
5.3.3	Custom Independents . . . . .	18
5.3.4	X Step . . . . .	18
5.3.5	Blob Detector . . . . .	18
5.3.6	Bar Charts and Histograms . . . . .	18
<b>6</b>	<b>Multiple Datasets</b>	<b>18</b>
<b>7</b>	<b>Handling Digitized Data</b>	<b>19</b>
7.1	Sort Data . . . . .	19
7.2	Formatting Dates . . . . .	19
7.3	Number Formatting . . . . .	21
7.4	Export to .CSV File . . . . .	21
<b>8</b>	<b>Distance and Angle Measurements</b>	<b>22</b>
<b>9</b>	<b>Import and Export JSON</b>	<b>22</b>
<b>10</b>	<b>Custom Scripts</b>	<b>22</b>

# 1 Introduction

A large quantity of technical data is available only in the form of plots and images. In these images, it is easy to visualize the relationship between the variables involved, but recovering the exact numerical values of the data is usually a tedious and error prone process. To aid this time consuming task of data recovery, many digitization tools have been developed over the years, but this task remains daunting and prone to errors. The common complains with the existing tools are as follows:

- Limited in features, often supporting only a few specific plot types.
- Compatible with only some specific operating systems.
- Difficult to use or prone to errors.
- No longer maintained by the original developer.
- Closed source or no freedom to modify the code.

Because of the above limitations with the current digitizing tools, WebPlotDigitizer was developed to facilitate easy and accurate data extraction from a variety of plot types. This program has been built using HTML5 which allows it to run within most popular web browsers and does not require to be installed by the user (desktop versions are also available for offline usage). This is distributed free of charge as an opensource software and since its creation in 2011, this tool has gained thousands of users and has been cited in many published articles. A screenshot of a typical session of the software is shown in Figure 1.

## 1.1 History

WebPlotDigitizer was initially developed while working on my graduate studies at the University of Notre Dame. Having to extract data from many publications for comparing and contrasting my own findings in the lab was a time consuming task. The search for a tool to aid this process usually ended in realizing that most of the existing tools for this purpose did not fulfill many of the requirements.

Some of the experimental work in the lab required me to learn some basic image processing techniques which eventually formed the basis of the automatic detection algorithms used here. Image processing knowledge along with some interest in learning the very popular HTML5 APIs were a perfect match to create a web based data extraction tool like this.

WebPlotDigitizer is now used by thousands everyday and because of this interest, I have continued to maintain and improve the software in my spare time.

## 1.2 User Manual and Tutorials

This user manual describes the various capabilities of the software and aims to help the user in making an effective use of the software. This manual may be updated continuously to

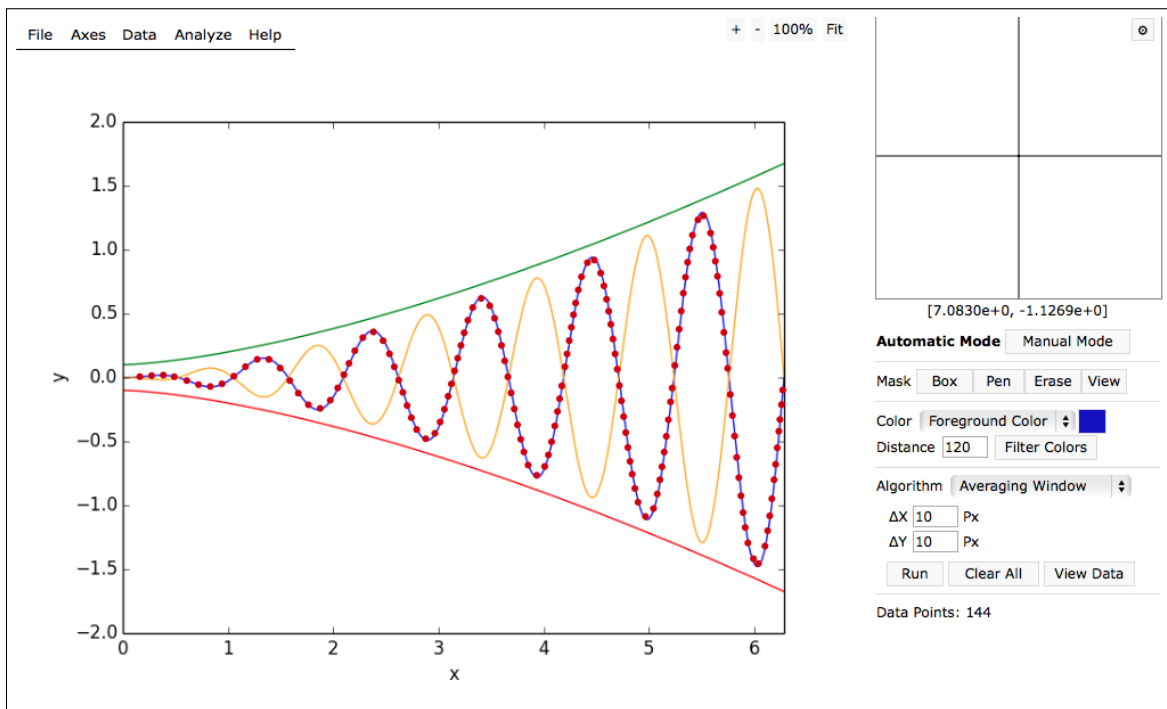


Figure 1: Screenshot of WebPlotDigitizer showing the data points recovered on a plot via automatic detection.

match the latest deployed version of the software. A few video tutorials are also available at <https://automeris.io/WebPlotDigitizer>. If the you want me to document the underlying theory and algorithms, then please send me a request over email.

### 1.3 License

WebPlotDigitizer is distributed under the GNU Affero General Public License version 3 by Ankit Rohatgi. For the complete terms and conditions, please refer to <https://www.gnu.org/licenses/agpl-3.0.html>

### 1.4 Source Code

WebPlotDigitizer is an open source software (see above). The source code can be obtained from GitHub (<https://github.com/ankitrohatgi/WebPlotDigitizer/>).<sup>1</sup> Desktop versions are also available for Windows, Linux and Mac OS on the website <https://automeris.io/WebPlotDigitizer/download.html>. Feel free to contact via email if you wish to contribute to this project.

### 1.5 Availability

The latest stable version of the software can be used directly from the website <https://automeris.io/WebPlotDigitizer>.

### 1.6 Supported Browsers

Online version of WebPlotDigitizer should be compatible with most recent versions of popular browsers such as Google Chrome, Firefox and Safari. Internet Explorer and Microsoft Edge are not tested.

### 1.7 Citing WebPlotDigitizer

If you wish to cite WebPlotDigitizer in any of your works, then please use the following information:

Author	Ankit Rohatgi
Title	WebPlotDigitizer
Website	<a href="https://automeris.io/WebPlotDigitizer">https://automeris.io/WebPlotDigitizer</a>
Version	4.6
Date	September, 2022
E-Mail	ankitrohatgi@hotmail.com
Location	Pacifica, CA, USA

---

<sup>1</sup>For an efficient management of e.g., the illustrations of the documentation, this project uses git's Large File Storage (<https://git-lfs.github.com/>). For a successful local checkout of the repository, this typically requires an additional package (example Debian, <https://tracker.debian.org/pkg/git-lfs>) to the installation of git itself.

## 1.8 Reporting Issues

Feel free to contact via e-mail to report issues or offer suggestions. If you are comfortable with GitHub, then please use the issue tracker for this project: <https://github.com/ankitrohatgi/WebPlotDigitizer/issues>.

## 1.9 Data Privacy

WebPlotDigitizer's image analysis code runs entirely on the user's computer and does not store the loaded images or data on to any server. When *Graph in Plotly* option is selected, the digitized data is transmitted to Plotly (<https://plot.ly>) servers. For a detailed privacy policy, please refer to <https://automeris.io/WebPlotDigitizer/privacy.html>.

## 1.10 Funding

WebPlotDigitizer is not funded by any organization. This project is entirely supported by the PayPal donations from generous supporters and my personal time and effort.

# 2 Loading Plots

The image file containing the figure to be analyzed can be loaded into the software in the following ways:

1. **Drag & Drop Operation:** Image can be dragged and dropped from the file browser on to the image viewing area of the application.
2. **File Menu → Load Image:** Browse for a file on the hard disk to load.
3. **Copy-Paste from Clipboard:** This is only supported in Google Chrome web browser. An image selected by copying in a PDF or an image viewer can be pasted on to the software via a simple copy-paste operation.
4. **File Menu → Webcam Capture:** A snapshot taken from the webcam can also be used. For best results, the webcam should be pointed directly along the normal to the plot surface. In the future, some perspective transformation tools might be added to WebPlotDigitizer to compensate for the distortions when the camera is not perfectly aligned.

## 2.1 Supported Image Formats

WebPlotDigitizer relies on the image formats supported by the HTML5 *canvas* element. Most browsers support common image formats such as JPEG, PNG, BMP and GIF. Since the support for an image format depends on the browser used to access the software, please refer to your browser's manual for details. For popular browsers, you can also refer to Wikipedia ([https://en.wikipedia.org/wiki/Comparison\\_of\\_web\\_browsers#Image\\_format\\_support](https://en.wikipedia.org/wiki/Comparison_of_web_browsers#Image_format_support)).

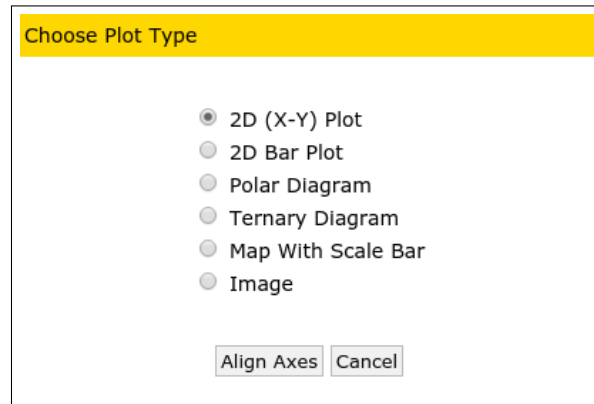


Figure 2: Popup with plot types that are supported in the software.

### 3 Calibrate Axes

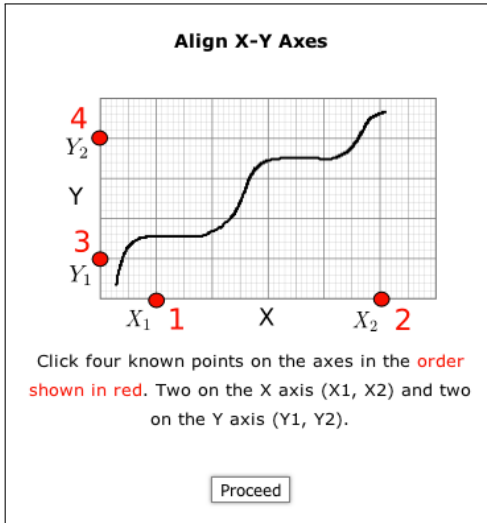
After loading the desired image, you should specify the type of axes that is used in the plot. This step is required for the software to correctly map the image pixels to the corresponding data values in the image. Depending on the plot type, you will have to select a few known points on the axes. On clicking the *Axes*  $\rightarrow$  *Calibrate Axes* menu item, you should be presented with the menu shown in Figure 2.

#### 3.1 2D (X-Y) Plot

Most plots that are on a two dimensional cartesian coordinate system fall under this category. Plots with non-orthogonal axes or plots rotated by an angle should work just fine, but images with perspective distortions will not yield accurate results. You can use image editing tools such as Gimp (<https://www.gimp.org/>) to correct the perspective distortions before using WebPlotDigitizer.

After selecting the 2D (X-Y) plot option, another popup window appears with instructions to click on two points on the horizontal axis and two points on the vertical axis. After reading these instructions, and closing this popup, click on two points on one of the axes ( $x_1, x_2$ ) and two on the other ( $y_1, y_2$ ). That is, you are free to attribute ( $x_1, x_2$ ) to either abscissa, or ordinate in first place. For better accuracy during the digitization process, pick the points that are as far away from each other as possible. Note down the data values at ( $x_1, x_2$ ) and ( $y_1, y_2$ ) positions as you will be required to enter those once the four points have been clicked on.

During selection of the four points, you can also make fine adjustments to the location of a calibration point that is highlighted in green using keyboard cursor keys. Pressing the *Shift*-key along with the cursor keys will move the point at a faster rate. To highlight a different point than what is already selected, you just need to click on the other point after the four points have been placed on the screen. Zooming into the image can help improve



(a) Mark four points to align axes

**X and Y Axes Calibration**

Enter X-values of the two points clicked on X-axis and Y-values of the two points clicked on Y-axes

	Point 1	Point 2	Log Scale
X-Axis:	0 <input type="text" value=""/>	1 <input type="text" value=""/>	<input type="checkbox"/>
Y-Axis:	0 <input type="text" value=""/>	1 <input type="text" value=""/>	<input type="checkbox"/>

\*For dates, use yyyy/mm/dd format (e.g. 2013/10/23 or 2013/10). For exponents, enter values as 1e-3 for 10<sup>-3</sup>.

OK

(b) Specify data values at the selected points to calibrate the axes.

Figure 3: Calibrating a 2D (X-Y) plot.

the accuracy.

After the four points have been clicked on, press *Continue* and another popup window will appear where you will be required to enter the values at these points. This helps the software map the image pixels corresponding to data points to their actual values when the image is digitized. To view the mapping equations, select the *Transformation Equations* option from the *Axes* menu.

### 3.1.1 Calibration Value Format

Like most computer programs, WebPlotDigitizer accepts integers (e.g. 1, 2, 3 etc.) or floating point numbers (e.g. 3.14159). Some extra things to keep in mind are as follows:

1. Fractions are not computed as numbers. Entering 1/2 (for example), will not be considered as 0.5.
2. For exponentials, the caret symbol (^) is not recognized and the values have to be entered as 1.45e-10 for  $1.45 \times 10^{-10}$  (for example).
3. **Dates:** This is enabled only for 2D (X-Y) Plots. At the time of calibration, the dates have to be entered in the format shown below. With the final digitized data, however, results can be formatted in many different ways (see section 7.2).



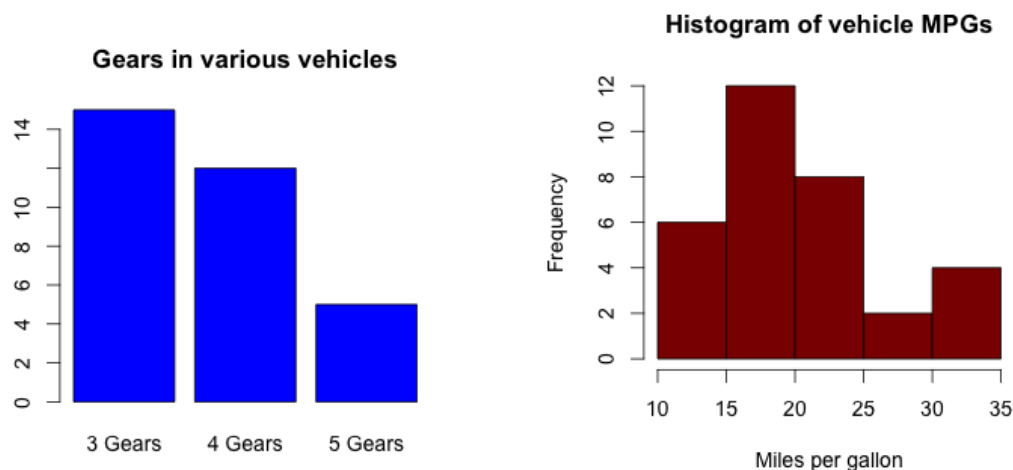
Table 1: Input format for dates in a 2D (X-Y) Plot.

Date	Format	Examples
Year, Month, Date and Time	yyyy/mm/dd hh:ii:ss	2012/10/23 05:11:06.25
Year, Month, Date and Hour	yyyy/mm/dd hh	2012/10/23 05
Year, Month and Date	yyyy/mm/dd	2012/10/23, 2012/10/5 or 2012/10/05
Year, Month	yyyy/mm	2012/10 or 1989/5
Year	yyyy	2012 (treated as any integer)

## 3.2 Bar Charts

### 3.2.1 Bar Charts vs. Histograms

For bar charts with a continuous variation of data along the data axis, but with only a discrete set of labels on the other, select the *2D Bar Plot* option when calibrating the axes. For plots such as histograms where both the axes are continuously varying, you should still select the *2D (X-Y) Plot* option and use the manual mode or the *histogram* algorithm for data extraction (see Figure 4).



(a) For bar charts with discrete items on the independent axis, use the 2D Bar Chart option when calibrating the axes.

(b) For histograms, both axes have continuously varying quantities. For this case, use the 2D (X-Y) Plot option and select the *histogram* algorithm.

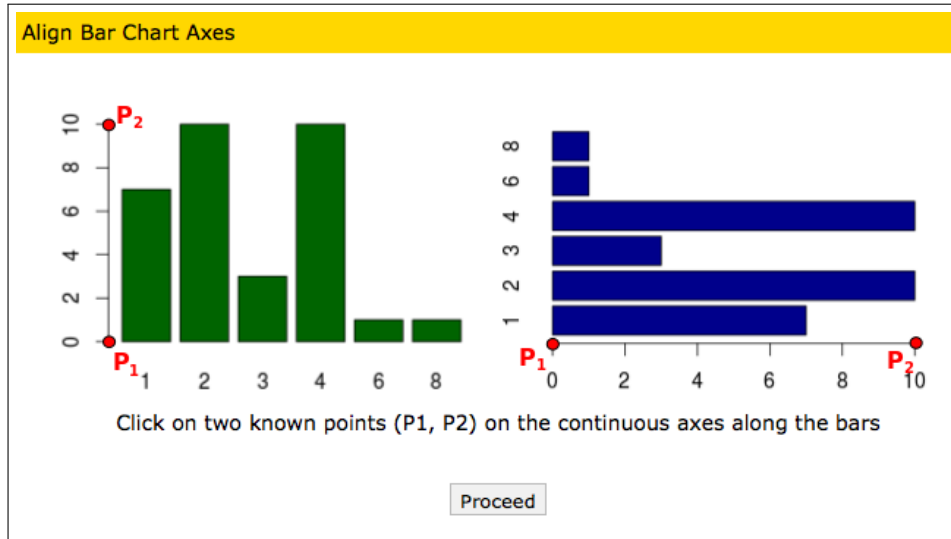
Figure 4: Bar Charts vs. Histograms

### 3.2.2 Axes Calibration for Bar Charts

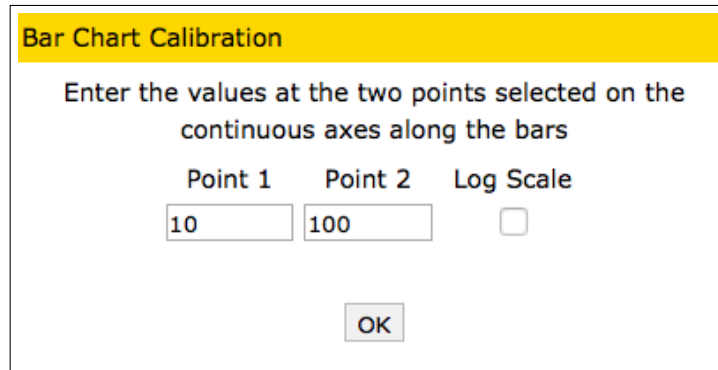
Select the *2D Bar Plot* option when calibrating the axes and follow the instructions to select two points on the continuously varying data axis that is along the bars. Next, specify the

values at these two points so that WebPlotDigitizer can construct a mathematical relationship between the pixel position along this axis and the corresponding data value (see Figure 5).

After calibration, you can manually click on the bars to mark the points or use the automatic extraction mode where you can use the color of the bars to extract data points. For bar charts, you can also edit the labels of each individual data point.



(a) Instructions for calibrating bar chart axes.



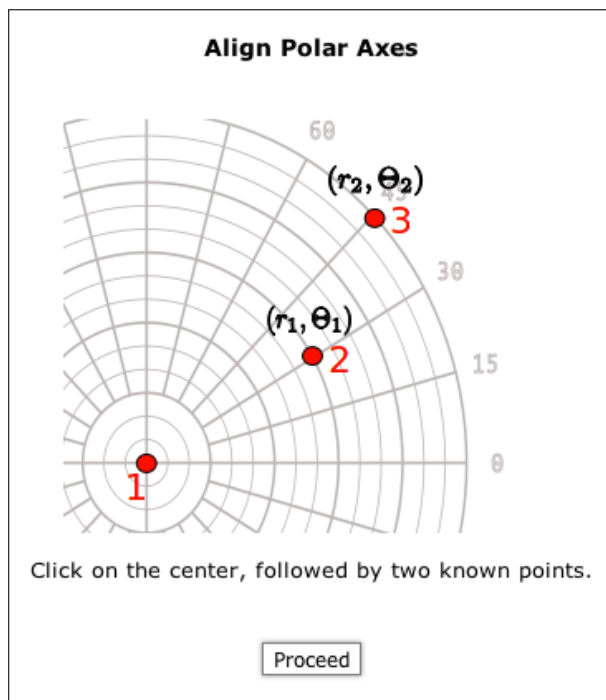
(b) Specify values at the two points selected on the continuous axis.

Figure 5: Calibrating axes for 2D Bar Plots.

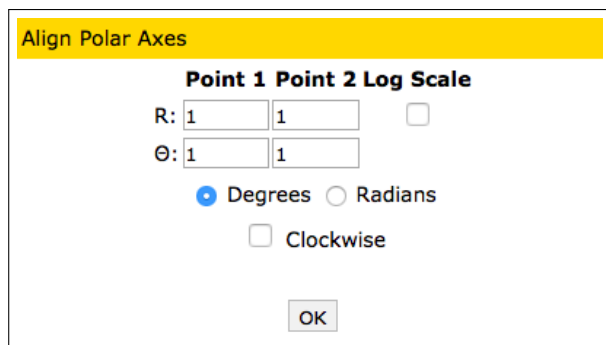
### 3.3 Polar Diagram

Select this option if the data points in the image are plotted on a polar axes. On selecting this, you will be required to click on three known points including the center of the polar diagram (Figure 6). After clicking on 3 points, you can also select the axes orientation and select Degrees or Radians for the angle. For log-polar diagrams, you can select the log scale

option. The values entered here also have to follow the format similar to 2D (X-Y) Plots. Dates can not be used for values in this case.



(a) Mark three points to align axes



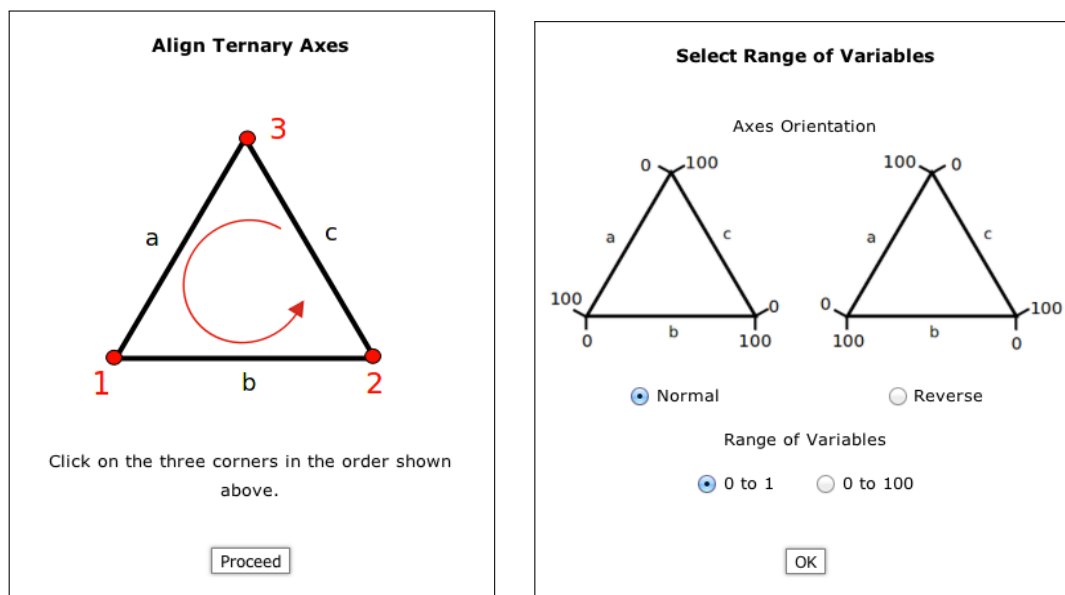
(b) Specify values at the two points other than the origin.

Figure 6: Alignment for Polar Diagram.

### 3.4 Ternary Diagram

Ternary phase diagrams are harder to interpret than simple two dimensional cartesian or polar plots. Using this software to recover data makes the process of data recovery extremely straightforward and thus reduces the possibility of misinterpreting the data. For this type of

plot, simply mark the three corners as shown in the instructions and then specify the range of variables and orientation of the diagram (Figure 7).



(a) Mark three points to align axes

(b) Specify values

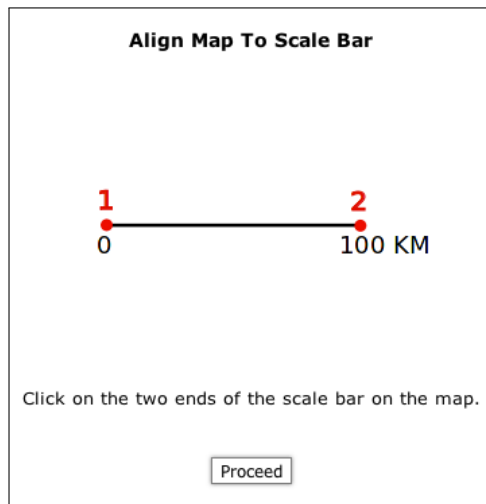
Figure 7: Alignment for Ternary Diagram.

### 3.5 Maps/Microscope Images

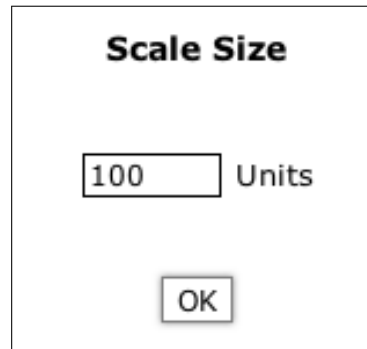
The *Map with Scale Bar* option is similar to 2D (X-Y) Plots and is to be used for images that only have scale information (e.g. microscope images or maps). To calibrate the pixels to the scale bar in these images, simply click on the two ends of the scale bar and enter the scale value. You can also specify the label to be used for the unit (Figure 8). The coordinates reported by the software assume the origin to be located at the top left of the image with positive  $y$ -axes pointing downwards. The  $(x, y)$  values that are generated are scaled using the value entered during calibration.

### 3.6 Image (Align to Image Pixels)

This plot type is similar to the Map plot type and is provided just as a convenience to work with generic images. If you select this type, no calibration information is required as the software will just do a 1:1 mapping from the image pixels to data values. This plot type is useful when the exact pixel locations of the image features are required.



(a) Mark ends of the scale bar



(b) Specify the scale value and the unit label

Figure 8: Alignment for maps and microscope images.

## 4 Grid Removal

The automatic extraction algorithms of WebPlotDigitizer rely on the color differences between the data points or curves and the background. This approach works only when there are very few background artifacts of the same color as the data. In many plots with grid lines, it is difficult for the extraction algorithms to distinguish between the background grid lines and the data curves (often in the same color). For such plots, a grid removal tool has been added, that can be used to remove the interfering horizontal and vertical lines present in the image before data extraction.

To perform a grid removal operation, select the *Remove Grid* option from the *Axes* menu. A new set of controls should appear on the side as shown in Figure 9.

- **Mask Editing:** Mark a rectangular region on the image after clicking the *Box* button. This will specify the region in which grid removal is to be performed. Use the *Clear* button to delete the selection. *View* button will simply display the box selection.
- **Color Selection:** Use the *Pick* button and specify the background color of the plot. If the *Background Mode* option is unchecked, then the selected color should be the foreground color of the grid lines. The number next to the *Pick* button is the acceptable distance from the selected color. With some hit-and-trial, you pick the distance that works the best. The *Test* button can be used to test the color detection.
- **Horizontal and Vertical Parameters:** By default both horizontal and vertical grid lines will be removed. To turn off detection in one of these directions, just uncheck the appropriate option. The *X%* and *Y%* parameters are used to specify the percentage

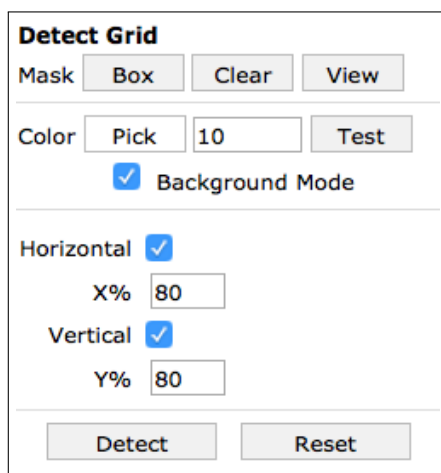


Figure 9: Grid removal controls.

of the height or the width that is taken up by the grid line. For solid grid lines, this number will be higher. For dotted or dashed lines, this number will be smaller.

To run the grid line removal tool, press the *Detect* button. If the results are not satisfactory, then you can adjust the various parameters and click *Detect* again. Hit *Reset* to restore the original image.

## 5 Acquire Data

Once the plot axes have been calibrated, you can begin selecting data points on the image. Also note that the numbers below the zoom window reflect actual data coordinates corresponding to your mouse position on the image. If you see incorrect numbers here, then perhaps incorrect calibration values were entered. You must repeat the axes calibration in this situation.

WebPlotDigitizer should also show a side panel with the data acquisition controls (Figure 10) when the axes are aligned. This sidebar can also be brought up by selecting the *Acquire Data* option in the *Data* menu. The data acquisition can be performed in either manual or automatic mode. You can alternate between the two modes at any time. In the manual mode you can add, adjust or remove data points by manually clicking at the desired locations. In the automatic mode, you can set up and execute an extraction algorithm that can differentiate between the data points and the image background and identify several data points in a short time.

### 5.1 Manual Mode

The controls available in this mode are as follows:

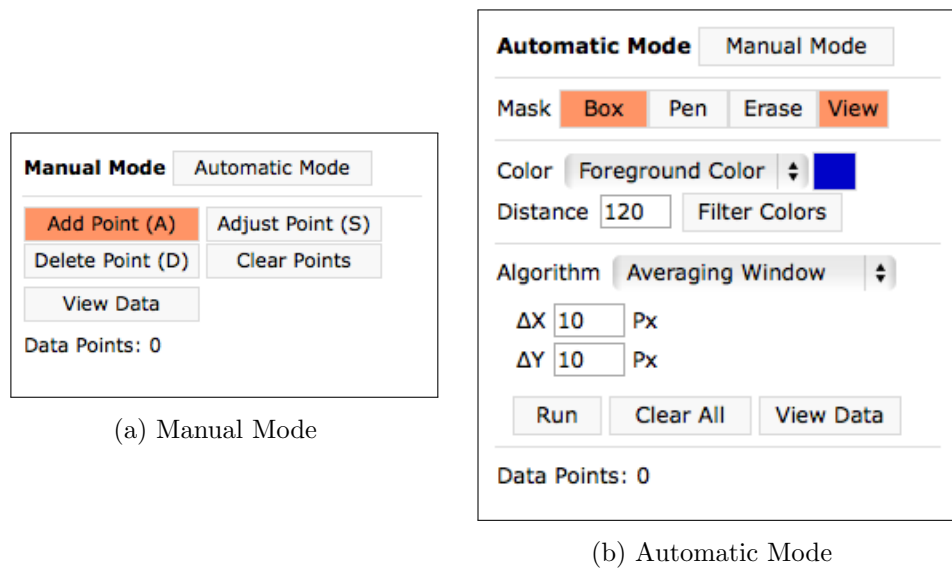


Figure 10: Data acquisition controls.

1. **Dataset:** Pick the dataset from the list to specify the data series the data points are to be added to (or modified). See section 6 for further details.
2. **Automatic Mode:** Switch to automatic extraction mode.
3. **Add Point (A):** After clicking this button or pressing the 'A' key, you can click in the image area to add data points. Any point added on the image will automatically be converted from on-screen pixels to data values utilizing the axes calibration.
4. **Adjust Point (S):** After clicking this or pressing the 'S' key, you can click on an existing data point to select it. The selected point can then be repositioned using the cursor keys on the keyboard. Press Shift+Cursor key for a faster rate of movement.
5. **Delete Point (D):** After clicking this button or pressing the 'D' key, you can click on a previously added data point to delete it.
6. **Clear Points:** This deletes all data points added on the image. This does not clear the axes calibration.
7. **View Data:** This launches a popup dialog where the collected data can be viewed, exported to a CSV file or Plotly. See section 7 for details.
8. **Edit Labels:** This option is only available for 2D Bar Plots. After clicking this button, you can edit the labels for a data points by clicking on them.

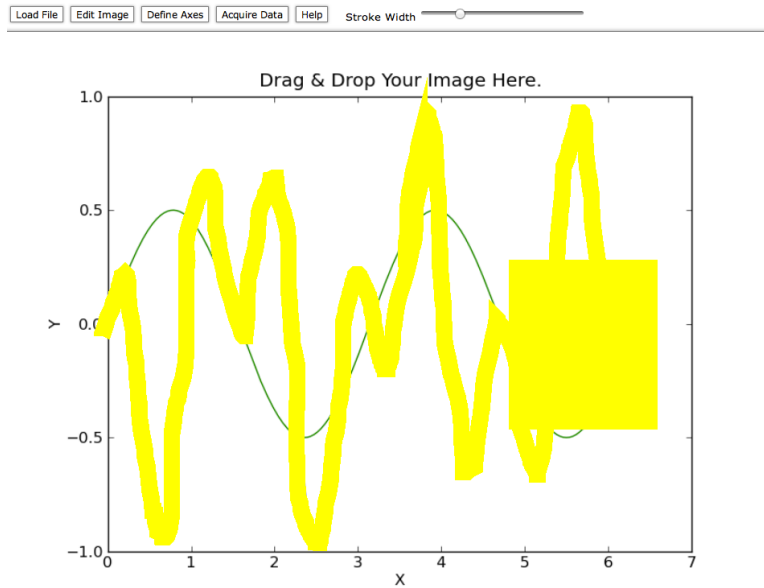


Figure 11: Use the *Box*, *Pen* and *Erase* tools to mark the region containing the required data (automatic extraction).

## 5.2 Automatic Mode

The controls available in the automatic mode are used for selecting the appropriate algorithm and providing the necessary inputs required for automated extraction of data points.

Automatic data extraction relies on separating the color of the data points or curves from the background in the image. The extraction algorithms can work in two modes of color extraction: Foreground mode and Background mode. In the foreground mode, the algorithms look for the foreground color specified for the data and ignore everything else. In the background mode, the algorithms include everything except the background color as potential data points. If the data points or curves of interest are uniformly colored (approximately), then the foreground mode may be more suitable. Otherwise if the background is uniformly colored (approximately) and the curve or data points are not then the background mode may be more suitable.

The extraction algorithms also need to know the region of the image to be searched for the specified colors. The software does not search the entire image as in many cases the data point or the curve colors may be present in non-data parts of the image as well. To specify the region of interest, use the *Box*, *Pen* and *Erase* tools (described below) to paint over a yellow mask over the data part of the image as shown in Figure 11. The extraction algorithms will look for data only under the yellow colored region. The various controls available in this mode are as follows:

1. **Manual Mode:** Switch to manual mode. You can switch between these modes at any



time.

2. **Dataset:** Pick the dataset from the list to specify the data series the data points are to be added to (or modified). See section 6 for further details.
3. **Mask Controls:** These controls are to be used to mark the search region for the extraction algorithms. Use the following drawing tools to highlight the approximate region containing data. You should try to exclude all features that are similar in color to the desired data points.
  - (a) **Box:** This is used to mark a rectangular region to be used during the search for data points.
  - (b) **Pen:** This is used for marking areas of the image to be included in the search for data points via free hand drawing.
  - (c) **Erase:** This is used to unmark the areas marked using the *Box* or *Pen* tools.
  - (d) **View:** This is used to simply view the marked search region.
4. **Color Controls:** Use these controls to specify the color of the data points and an approximate RGB distance from this color that is acceptable. You either need to specify the foreground color of the data or the color of the background. In most cases, the distance value will require some hit-and-trial based on the results generated by the extraction algorithm.
  - (a) **Foreground/Background Color Selection:** Select the appropriate option from the drop-down menu and click on the colored button on the right to show a popup for color selection. The popup dialog also shows a list of most dominant colors at the bottom that can be selected.
  - (b) **Distance:** Change this value to adjust the range of colors around the selected color to be accepted as foreground or background.
  - (c) **Filter Colors:** Test the color selection and the specified distance value. The area highlighted in yellow is the region that will finally be used by the automatic extraction algorithms.
5. **Algorithm Selection:** Select the appropriate algorithm depending on the application. Each algorithm is described in detail the next section.
6. **Run:** Start the auto-detection algorithm. After this is completed, the detected points should appear over the image. If necessary, adjust the parameters of the extraction algorithm, mask or color settings and run the algorithm until you are satisfied. You can also switch to manual mode and edit these data points.
7. **Clear All:** Clear all data points in the current dataset.
8. **View Data:** This launches a popup dialog where the data can be viewed and exported to a CSV file or Plotly. See section 7 for details.

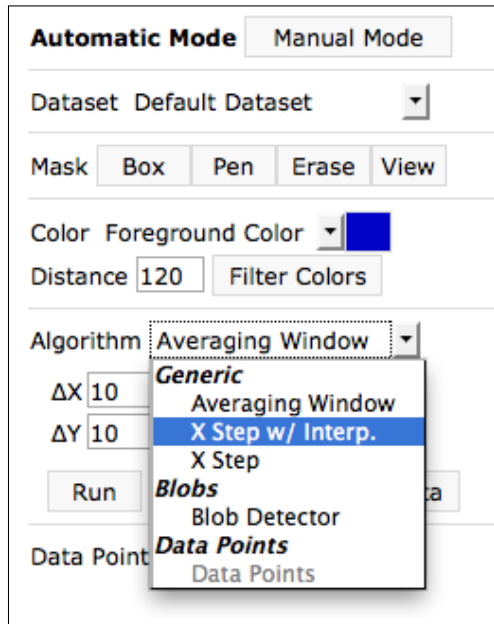


Figure 12: Four automatic extraction algorithms are available. These are subject to change in the future releases.

### 5.3 Digitization Algorithms

Four different digitization algorithms are available in WebPlotDigitizer as shown in Figure 12. The *Averaging Window* algorithm is set as the default as it is usually suitable in most simple cases.

#### 5.3.1 Averaging Window

As mentioned above, this is probably the most useful algorithm and is useful across many plot types. The data points detected by this algorithm are determined by calculating the average locations of pixels with acceptable color that lie in small regions that are  $\Delta X$  pixels (on-screen) wide and  $\Delta Y$  pixels (on-screen) tall. As a user, you should increase the size of this window for thick lines or large data points and decrease it for thin lines. If you see multiple points incorrectly detected across the width of a thick data curve, then you need to increase the numbers specified here. The fact that this requires on-screen pixels may be changed in the future so that the values in actual units in the current axes can be specified.

#### 5.3.2 X Step with Interpolation

This algorithm is available only for non log-scale 2D (X-Y) axes plots. In the future, this will be extended to other axes types. This algorithm can identify data points at regular intervals on the  $x$ -axis that fall between  $X_{\min}$  and  $X_{\max}$  and  $Y_{\min}$  and  $Y_{\max}$ . The data points are spaced at an interval  $\Delta X$  units apart. This algorithm interpolates over missing data

using cubic splines and is therefore suitable even for curves with dotted lines or a series with just data points. The *Smoothing* value can be increased from zero (for example, try 0.5) to average over a larger neighborhood around the data points. This is useful for reducing noise in the captured data. (Also see: [https://automeris.io/WebPlotDigitizer/blog/posts/discontinuous\\_data.html](https://automeris.io/WebPlotDigitizer/blog/posts/discontinuous_data.html))

### 5.3.3 Custom Independents

This algorithm is only available for 2D (X-Y) axes plots (without dates) for now. This allows the user to specify custom X values at which automatic digitization should be attempted. Use the *Curve Width* option to account for thicker curves (or noisy data points).

### 5.3.4 X Step

This is a simplified version of the *X Step with Interpolation* algorithm. This does not use cubic splines to interpolate over data. The *Line Width* parameter is the *y*-direction thickness of the curve in on-screen pixels (In future, this will be updated to use the appropriate units in Y).

### 5.3.5 Blob Detector

This algorithm is useful for counting number of objects, determining the location of their centroids, calculating areas and also first moments of continuous objects. All objects that are within the *Min Diameter* and *Max Diameter* range are added as data points. (Also see: [https://automeris.io/WebPlotDigitizer/blog/posts/blob\\_detection.html](https://automeris.io/WebPlotDigitizer/blog/posts/blob_detection.html))

### 5.3.6 Bar Charts and Histograms

## 6 Multiple Datasets

For plots that have multiple sets of data, you can capture data for one dataset, export the CSV file, then clear all the data points and start over to collect data for the next set. This may be acceptable in many cases, but this is can be somewhat inconvenient if you wish to return to an old dataset and modify some data points. Also, this way, you will not be able to store data for all the datasets in an exported JSON file (see section 7).

A better option is to create a separate dataset for each series you wish to capture by selecting the *Add* option in the *Manage Dataset* popup that can be accessed from the *Data* menu (Figure 13). From this popup, you can also delete or rename a dataset. After adding multiple sets, you can then switch between these from the dataset menus available in the manual and automatic extraction controls and also in the *View Data* window. By default, one dataset called "Default Dataset" is available.

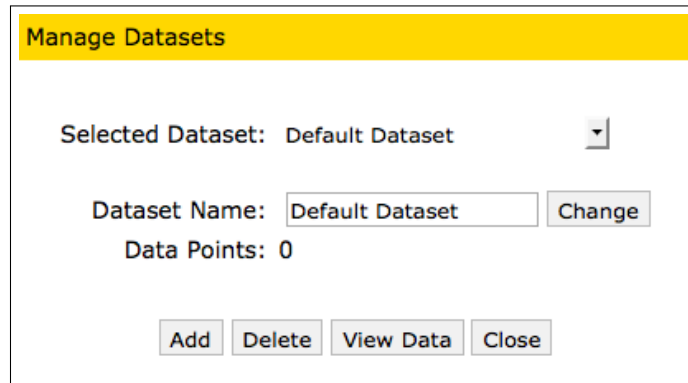


Figure 13: Add, delete or rename datasets corresponding to each series of data in the plot.

## 7 Handling Digitized Data

Once the required data points are marked on the image using the manual mode, automatic mode or a combination of both, the digitized values can be seen by clicking the *View Data* button. This presents a popup window as shown in Figure 14. Here, the digitized values can be sorted by the variable or in order of the distance between the points (Nearest Neighbor). The values can also be copied and used in common data analysis softwares. Recently, an option to send these values over to another cloud based data analysis and graphing software called Plotly (<https://plot.ly>) has also been added.

### 7.1 Sort Data

The digitized data can be left unsorted (Raw Output) or by one of the axes variables in ascending or descending order. The Nearest Neighbor option sorts the data depending on the distance of the points from each other.

### 7.2 Formatting Dates

If one or both of the axes in a 2D (X-Y) plot contain dates then fields to specify the output format of the values are also shown (Figure 15). In these fields, the following pieces of text are replaced with the corresponding part of the date to format the text (case insensitive):

Acquired Data

Dataset: blue curve ▾

Variables: X, Y

0.15708,	0.01801
0.26857,	0.02620
0.38007,	0.02984
0.48396,	0.01708
0.60306,	-0.01333
0.71455,	-0.04497
0.82604,	-0.05953
0.93754,	-0.03883
1.05005,	0.02067
1.16053,	0.09474
1.27202,	0.15223
1.38351,	0.15170
1.48994,	0.08347
1.58116,	-0.01470
1.66225,	-0.11285
1.75854,	-0.20382
1.87003,	-0.23242
1.96632,	-0.16855

Select All Download .CSV Graph in Plotly\* Close

\*Plotly is a secure data analysis and graphing site with data sharing and access controls.  
Visit <http://plot.ly> for details.

**Sort**  
Sort by: Raw ▾  
Order: Ascending ▾

**Format**  
Number Formatting:  
Digits: 5 ▾ Fixed ▾  
Ignore  
Fixed  
Precision  
Exponential  
Format

Figure 14: CSV formatted digitized data. This can be sorted, pasted to a .CSV file or graphed in Plotly.

Sort by: Raw Output ▾ Order: Ascending ▾

Date Formatting: X mmm yy ▾ Format

Variables: x, y

Figure 15: Field to specify formatting of dates. This appears on X, Y or both axes depending on which variables contained dates at the time of axes calibration.

Text	Replaced With	Example
yyyy	Year, all digits	2012
yy	Year, last two digits	98 for 1998
mmmm	Month, full name	January
mmm	Month, short name	Jan
mm	Month, numeric	10
dd	Date	23
hh	Hour	05
ii	Minutes	12
ss	Seconds	5
ss.frac	Fractional Seconds	5.25

A few examples are shown below:

Format Field Text	Date Shown (Example)
dd-mm-yyyy	23-10-2012
mmm-yyyy	Oct-2012
yyyy abc mmmm	2012 abc October
mmm 'yy	Oct '12

### 7.3 Number Formatting

The number formatting of the digitized data columns can be changed by specifying the number of digits and the formatting style from the menu. Three formatting styles are available at the moment:

1. **Fixed:** Use this to display the desired number of digits after the decimal.
2. **Precision:** Use this to set the number of significant digits.
3. **Exponential:** Use scientific notation with the desired number of digits after the decimal.

The *Ignore* option disables the number formatting.

### 7.4 Export to .CSV File

Comma-Separated Values (CSV) format files are simple text files containing tabular data. Each row of line corresponds to a table row and the column values are separated by a character or a string (usually just a comma)<sup>2</sup>. Due to its simplicity, CSV format is supported in most data analysis softwares like Microsoft Excel, Matlab etc.

To save the values from WebPlotDigitizer into a CSV file, all you have to do is open your favorite text editor (e.g. Notepad on Microsoft Windows) and copy-paste the values shown

---

<sup>2</sup>For more information on CSV format, check out [https://en.wikipedia.org/wiki/Comma-separated\\_values](https://en.wikipedia.org/wiki/Comma-separated_values)

in the results popup window. Save the file with a .csv file extension. You can now use this file in your favorite data analysis software.

## 8 Distance and Angle Measurements

Besides extracting data points from plots, WebPlotDigitizer can also be used to make some simple distance and angle measurements on images. In the future, many more advanced analysis tools will be added. To make distance or angle measurements, choose the appropriate option from the *Analysis* menu.

For distance measurements, if the chosen plot type is a *Map* axes, then the distance is converted to the units specified during calibration. Otherwise distance values are shown in pixels on the original image (not same as on-screen pixels!). Angle measurements are not dependent on any axes. In each case, you can make a single measurement after pressing the *Add* button (or pressing 'A' key) and then clicking on the image. After the measurement is made, you can adjust the highlighted (green) point using the cursor keys on the keyboard (use Shift for faster movement). To change the highlighted point, click on the desired point. A measurement can be deleted after clicking the *Delete* button or pressing the key 'D'. *Clear All* will delete all measurement data in that mode. *View Data* will show a popup with all measurements that can be exported to *Plotly* or a *CSV* file.

## 9 Import and Export JSON

If you wish to save the axes calibration, data points in all the datasets and the measurements to a file for future use, then you can select the *Export JSON* option in the *File* menu. Note that this file does not contain the image and so you will also have to keep the plot image along with this JSON file. To use the information stored in the JSON file, simply use the *Import JSON* option in the *File* menu. The Export and Import feature can also be used to reuse axes calibrations across multiple plot images which are plotted on the same axes.

## 10 Custom Scripts

Users can also create custom scripts to automate some of the digitization tasks or add new functionality. The custom scripts have access to most of the WebPlotDigitizer's API and so it is possible to accomplish fairly complicated tasks. A few simple examples can be found on [https://github.com/ankitrohatgi/WebPlotDigitizer/tree/master/script\\_examples](https://github.com/ankitrohatgi/WebPlotDigitizer/tree/master/script_examples). If you need assistance with custom scripting, then feel free to contact me.