

Quanty Public Data

ReadMe

Quanty software and associated documentation and webpages is copyright of Jim Davies.
Public data offered by Quanty is in the public domain.

This document will help you to make sense of Quanty data. It is recommended that you try the Quanty game to be familiar with the context of Quanty's information collection.

Summary of Quanty Data

The Quanty game uses a set of image files. Each image is associated with 2 XML files:

- The annotation file, which contains information on where specific objects in the images are, and the coordinates used to draw outlines around them.
 - The Quanty data file, which refers to the objects in the annotations, and provides many answers to questions about them. The answers come from users playing the Quanty game who have been asked that question.
-

Quanty Uses LabelMe

The image files and annotation data (XML files that contain the information used to draw outlines around particular objects in each image during the game) come from the LabelMe project (<http://labelme.csail.mit.edu/>).

When you access Quanty image files and annotation files, you are accessing copies of a subset of the content available on LabelMe.

Quanty Game XML Data

Question Types

A Quanty question type can be either a measurement, where Quanty asks for a numerical input, or a Likert-scale question, where Quanty asks for a rating (ex: "How true is it that the man is inside the car"). These ratings are stored as integers from 1 (lowest) to 5 (highest).

Each question type is represented by a unique string. Here are the categories and the strings that fall into them.

Questions involving 1 object

Measurements:

- "height"
- "length"
- "width"
- "distance_fromCamera"
- "depth"
- "weight"
- "volume"
- "temperature"

Likerts:

- "height_L"
- "length_L"
- "width_L"
- "distance_fromCamera_L"
- "depth_L"
- "weight_L"
- "volume_L"
- "temperature_L"
- "beauty"
- "brightness"

Questions involving 2 objects

Measurements:

- "distance_between"

Likerts:

- "below"
- "above"
- "close"
- "far"
- "occluding"
- "inside"

Questions involving 3 objects

Measurements:

- (none)

Likerts:

- "between"

Objects

In each LabelMe annotation XML file, there are a series of <object> nodes. Each represents an object within the image. There is information like the name of the object, and the set of points

that form its outline.

In each Quanty data XML file, there is another set of <object> nodes (maximum 3) under each <question> node. These <object> nodes contain only a number. This is an index number referring to the <object> nodes in the LabelMe annotation file.

When the annotation XML is parsed, the set of <object> nodes reads as a list. The index numbers found in the Quanty XML are index numbers for this list.

EX: If the Quanty XML says
 <object>3</object>
...this refers to the third <object> node in the associated annotation XML.

Now, how can you translate a Quanty XML data <question> into an actual question?

When Quanty assembles a question involving multiple objects, it phrases that question using the objects in the order they are presented in the Quanty data XML. However, **Quanty also eliminates redundant questions.**

Example A: “How true is it that Object A is between Object B and Object C?” In this example, it would be unorganized to store information on another question that asks *Example B:* “How true is it that Object A is between Object C and Object B?” Notice that only the order of the objects has changed, but the question carries the same meaning. If Quanty generates this question, it will first make sure that Example A does not already exist in the database. If it does exist, Quanty will ask Example A instead, since the meaning is the same.

This extends to other question types where redundancy is possible, eg. “distance_between”

Values and Units

In Quanty data XML, the response that Quanty collects from users during the game is at the node `QuantyData > question > value`. There are at least two of these per question, and there may be many values if the question has been asked a lot.

If a user chose to pass the question, the content of the <value> node will read “Pass”.

If the user answered the question, the content of the node will be a number representing the user’s guess at a measurement. The unit of this measurement is indicated by the question type. All measurements are stored in what is known as the “base unit” of that question type. There are many available input units, but they are all converted to this single unit for storage.

Here is a categorization of question types, and their base units.

- Category: **distance** (base: **meters**)
 - “height”

- "length"
- "width"
- "distance_fromCamera"
- "depth"
- "distance_between"
- Category: **mass** (base: **grams**)
 - "weight"
- Category: **volume** (base: **litres**)
 - "volume"
- Category: **temperature** (base: **°C**)
 - "temperature"
- Category: **Likert (no unit)**
 - All other question types are Likert questions (multiple choice ratings) and are stored as integers from 1 to 5

“Stop Asking”

In the Quany data XML, the node `QuanyImage > question > stop_asking` is managed by Quany staff. It contains a boolean value, and when it is set to “true”, Quany will no longer ask this particular question during a game. Reasons for this include:

- The question has recieved enough data. If there are sufficient responses to calculate a meaningful average, then Quany game time is better spent asking other questions.
- The question has too many responses of “Pass”. If most users that encounter this question are passing on it, the question probably does not make much sense; after all, each question is randomly generated.

You may wish to keep this in mind when developing your own applications for Quany data.

Data Versions

On the Quany public data page (www.quanygame.com/data.html) you will find the following data available for download.

1. Images
2. LabelMe Annotation File
3. Quany Data - On Demand
4. Quany Data - Backups

You can download the Quany data files on demand, meaning they are the most up to date files generated for you on the fly, or you can download the Quany data files from one of the available backups. You do not need to download both and it is recommended you download all the files from the same backup date.