Truenumbers for Excel Cheat Sheet

Introduction to truenumber data

Computers have always played two roles:

- Assisting us with the natural language that we work, think and communicate with (*like word processing, email, document management*)
- Computing with data represented in formal structures (*like SQL, UML, XML*) .

The first is administrative, leaving the meaning of content for the human reader. The second involves paring meaning down to simple schemas.

Truenumbers is a third way that indexes data by descriptive phrases in a structured natural language, kind of like the index in a book, or library card catalog. This allows both computation, and rich human description. Think of it as document management for many tiny documents, each describing a single item of data

Truenumbers grammar

Allowed sentence structures: **<subject phrase> has <property phrase> = <value>** **<property phrase> of <subject phrase> is** (or **=**) **<value>**

Examples: **moon has diameter = 1079.4 +/- 0.7 miles** **diameter of moon is 1079 mi**

Phrases can be nouns, and also use adjectives, and possessives indicated by "of"

Example: **moon of earth has equatorial diameter = 1079.4 +/- 0.7 miles**

Phrase-paths: phrases are equivalent to paths (like URLs) where **:** represents adjective, and **/** represents possession (of)

Example: **moon of earth => earth/moon** **equatorial diameter => diameter:equatorial**

Tags are phrase paths, attached to a truenumber to provide context or additional information.

Values can be a number (with error and units), a string, or a phrase.

Examples of truenumbers (@ indicates a tag)

moon has diameter = 1,079.4 miles

older brother of Robert Kennedy is John Kennedy
@figure:political

John Kennedy has younger brother = Kennedy:Robert
@figure:political

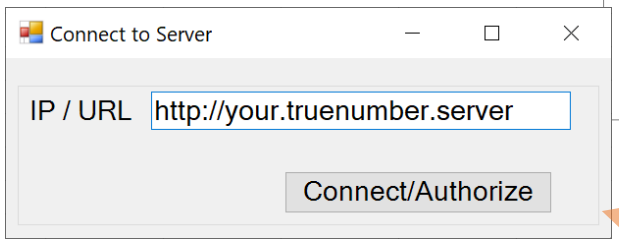
Melville/Moby-Dick has sentence:first = "Call me Ishmael"
@author:Melville:herman @date:publication:1851

The same truenumbers ed in Excel in the add-in's *standard outline format*

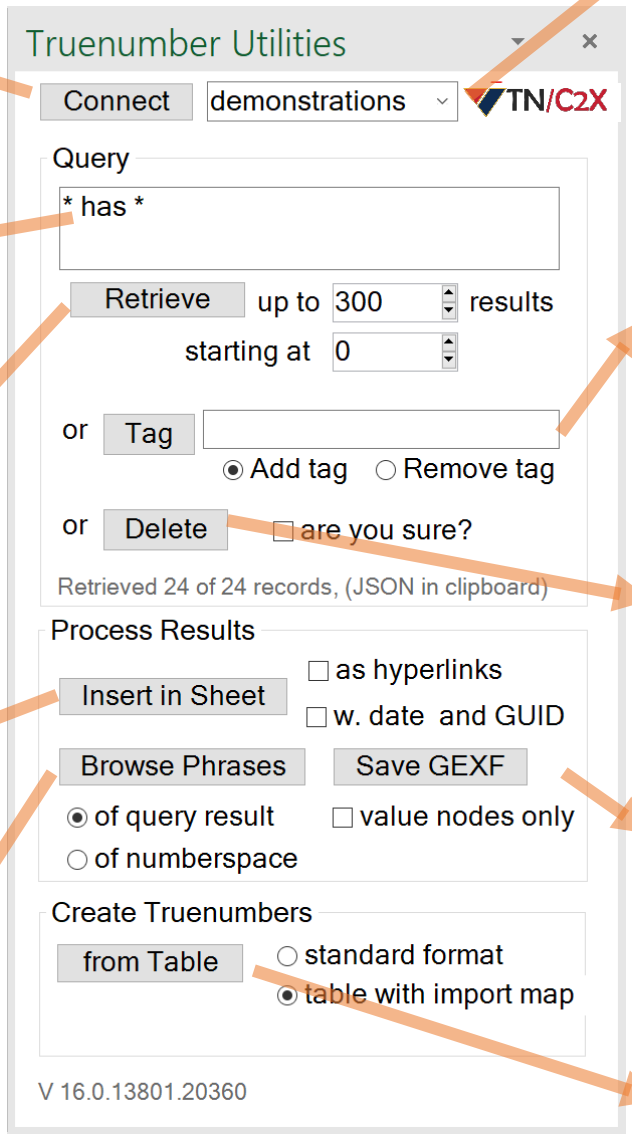
SUBJECT			PROPERTY		VALUE	UNITS	TAGS
Kennedy	John	has	younger brother	=	Kennedy:Robert		figure:political
	Robert	has	older brother	=	Kennedy:John		figure:political
Melville	/ Moby-Dick	has	first sentence	=	"Call me Ishmael"		author:Melville:herman date:publication:1851
moon		has	diameter	=	1079.4	miles	

Truenumbers for Excel Cheat Sheet

4/23/2021 11:34:37 AM



You must connect to a server by clicking the **Connect** button. Your administrator will have a default address there but you can change it



Once connected, this dropdown will allow you to choose one of the numberspaces on that server. Each numberspace is a separate collection of truenumber data. The default is **testspace** – consider it a scratchpad.

Here is where you type queries in the Truenumber Query Language. Queries are examples of truenumbers using wildcards. The query:

*** has *** -- will match all truenumbers

More about queries on the following pages.

The **Retrieve** button will downoad truenumbers that match the query, and store them within the add-in. Number boxes let you choose how many, and where to start in the sequence. It will report how many were retrived, and total matches in the numberspace.

The **Tag** button lets you add or remove tags from truenumbers in the numberspace matching the query w/o retrieving them, so this operation so it can apply to very large number of truenumbers. Type the tags in the box provided, comma-separated.

The **Delete** button will delete all truenumbers in the numberspace matching the query. No data is retrieved by this operation so it can apply to very large number of truenumbers. For safety, the are you sure? box must be checked first..

Insert in Sheet will format the truenumbers retrieved by query, and insert them in a standard format, one truenumber per row. Including the creation date and UUID of each is an option, off by default. You can also opt to have truenumber values be inserted as portable hyperlinks carrying the value's description onboard.

Save GEXF will export a file in gexf format suitable for display in compatible graph visualization software (see following pages)

Browse Phrases allows you to see the subject, property and tag taxonomies of the most recent retrieved query result, or of the entire numberspace. These can be viewed in list or tree form. The tree-view allows adding new taxonomy entries (details on pages following)

Create truenumbers **from Table** will read contents of the current worksheet, starting at the cursor position, in one of two formats (details to follow) and post the to the connected server numberspace.

Truenumbers for Excel Cheat Sheet

Create Truenumbers

from Table

◉ standard format

○ table with import map

Creating truenumbers from standard outline format

Rows in tables represent individual records of data, and each column, an aspect of the record. In this figure each row represents a county, with columns shown specifying the state, county and population. It's implicit that the state and county specify the subject of the row, and that the population is one of its properties, though from the data structure point of view, they are all just columns.

State	County	Population
Rhode-Island	Bristol	49,875
Rhode-Island	Kent	166,158
Rhode-Island	Newport	82,888
Rhode-Island	Providence	626,667
Rhode-Island	Washington	126,979
Hawaii	Kauai	72,293
Hawaii	Maui	167,417

Describing the highlighted cell to someone, we might say ***Kent county of Rhode Island has population = 166,158***. This sentence makes clear that Kent county is a part of Rhode Island, and has the property population, with a value of 166,158. In truenumber terms, every data *value* measures a *property* of a *subject*. We can show these relationships in a heirarchical form in Excel that we call *standard format*.

	A	B	C	D	E	F	G
1							
2	Rhode-Island	/county	Bristol	has	population	=	49,875
3			Kent	has	population	=	166,158
4			Newport	has	population	=	82,888
5			Providence	has	population	=	626,667
6			Washington	has	population	=	126,979
7							
8	Hawaii	/county	Kauai	has	population	=	72,293
9			Maui	has	population	=	167,417

In it, we've put the main subject of each value in the first column. In the next, the slash symbol / indicates that *county* is an aspect of *Rhode-Island*. The next column has no slash, so *Bristol*, for example, is taken as an adjective of *county*. In fact, if we "read" these columns right to left, and say "of" when we hit a slash, it reads "Bristol county of Rhode-Island".

So we have a kind of an outline form of the phrase. By using the word *has*, and an equal sign, we have something that combines a tabular form with a kind of sentence, and makes clear the distinction between subjects, properties and values. To add these to the numberspace as truenumbers, we would select cell A2, as shown, and click the **from Table** button with *standard format* indicated. If everythng was formatted properly, the system will report the number of posted truenumbers.

from Table ◉ standard format ○ table with import map

OK 7 truenumbers posted

Including units and tags on truenumbers

2	Rhode-Island	/county	Bristol	has	population	=	49.875	thousands	code:ISO-3166:US-RI

Standard format allows numeric values to have units of measure in the cell to the right. Above we see thousands specificied for population. Cells in subsequent columns can each contain a tag, expressed in path form. In this example, the tag path is equivalent to the phrase "US-RI ISO-3166 code" which gives us the state and country code as well as the ISO standard specifying it. There is no limit on the number of tags, extending out column by column.

Truenumbers for Excel Cheat Sheet

Create Truenumbers

from Table

○ standard format

◉ table with import map

Creating truenumbers from CSV data

CSV data in Excel appears as a table with its first row containing column headers. The Excel add-in can generate tagged truenumbers for the entire table using an *import map*, providing instructions for how to construct truenumbers and tags from data in each row. This feature has its limitations, so for some data it may be necessary to manipulate the data in Excel first, or to use a different import method like writing a custom script using the Truenumber API.

Put the *import map* in **row 1**. You will usually need to insert a blank row above the header. In each column of the import map, a special macro specifies what to do with that column. A blank ignores the column, and the row end *must* be marked with a cell containing "x".

	A	B	C	D	E	F
1	**sub3=$valt**	**sub2=/$valt**	**$sub\|mass\|$valn kg|solar-system**	**sub1=object:$valt**	**$sub\|description\|$vals|solar-system**	**x**
2	eName	type	mass_kg	orbits	description	
3	Moon	moon	7.35E+22	Earth	our moon	
4	Phobos	moon	1.06E+16	Mars	larger moon of Mars	
5	Jupiter	planet	1.90E+27	Sun	gas giant	

Import map row

Import macros

*Path variable*s - an arbitrary name ending in a number. For example, **myVar8** is a well-formed path variable. To reference a path variable, we precede the name with a dollar-sign as in **$myVar**. This concatenates all the myVar's in numeric order to build the value. If your segments don't include **:** or **/** separators, an **:** will be provided by default. You may also reference segments individually by including the number as **$myVar8** for example.

Built-in variables - **$valt** "tokenizes" the value in the current row,column, replacing spaces with "-" and so forth, to return a token. **$valn** expects a numerical value, and **$vals** returns a string, being whatever is in the cell, enclosed in double quotes. **#row** returns the number of the current row being processed.

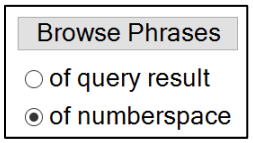
Example: columns A, B and D in row 4 would evaluate as **sub3= Phobos**, **sub2=/moon** and **sub1=object:Mars**, so **$sub** returns **object:Mars/moon:Phobos**

Truenumber specification – is 4 strings with variable references, separated by 3 vertical-bars. They represent **subject|property|value|tag1, tag2**. …

Example: columns C row 4, and E row 5, would produce the truenumbers **object:Mars/moon:Phobos has mass = 1.06 x 10^{16} kg @solar-system**, and **object:Sun/planet:Jupiter has description = "gas giant" @solar-system** respectively.

The General Idea

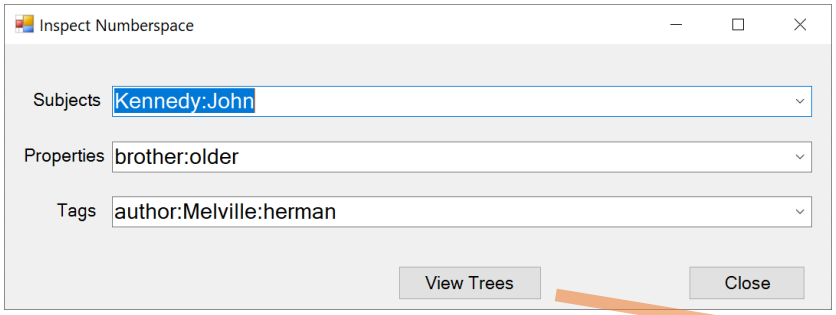
A row in a dataset implies a complex of information. The CSV import process intends to represent it more explicitly and naturally. In the example, we combine what the *orbit* and *type* columns tell us, and make it explicit in the phrase "**Phobos moon of Mars object**" or **object:Mars/moon:Phobos**. The CSV data doesn't explicitly say that it deals with objects in the solar system, so we add that information as a tag. There are no fixed rules for what information to put into subject phrases, properties, values or tags. So long as they are sensible and readable, it doesn't really matter. Users will be able to discover and search them.

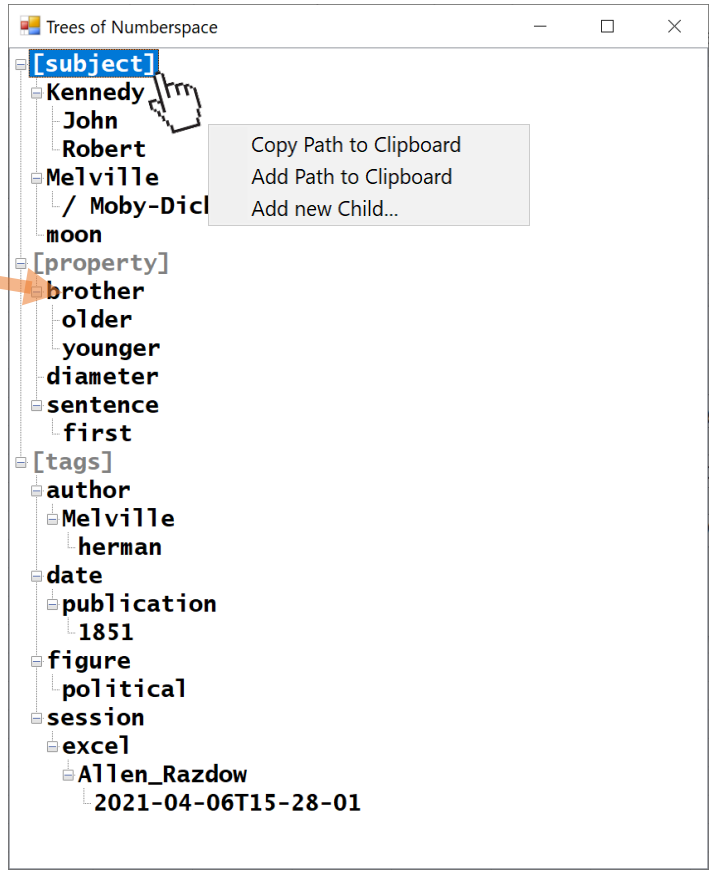


Browsing Phrase Vocabularies

To understand what's in your numberspace, or in the results of a query, it's helpful to browse the subjects, properties and tags in them. The Browse Phrases button allows this. It's set for the entire numberspace by default, which can take quite a few seconds for a big numberspace. After performing a query, the selection will change to browse the query result. The three dropdowns list all of the paths. To view the same information in tree form, click the View Trees button. Right-clicking on any tree node brings up a context menu with three commands:

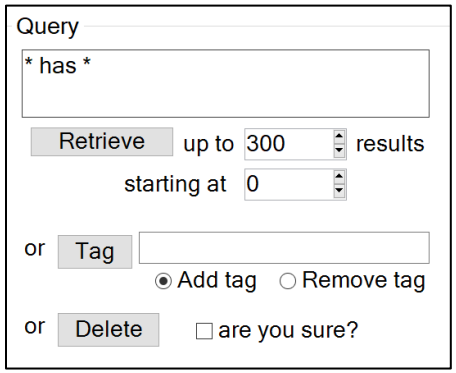
Copy Path to Clipboard puts the path to the selected node in the clipboard. **Add Path to Clipboard** adds the path to the selected node to the clipboard, allowing multiple paths to be copied. This is useful for quickly building queries in the query window of the add-in. Lastly, **Add New Child…** allows you to enter a name or path and add it as a child of the selected node. This lets you build taxonmies for later use. **NOTE** the "session" tag at the end of the tree. These are generated when you create new truenumbers from the add-in, containing the creation timestamp and MS-Office username of the Excel user. Session tags are NOT inserted into the worksheet when you use **Insert in Sheet**, but can be used in queries.





Truenumbers for Excel Cheat Sheet

Queries



Queries specify a set of truenumbers on the server. We can use the **Retrieve** button to pull them into memory in the add-in; or use the **Tag** button to add or remove tags from them on the server w/o retrieving. We can also (carefully) use the **Delete** button to remove them from the server.

*** has *** **will match every truenumber** ***NOTE: Queries are not case sensitive***

Patterns for subject or property phrase paths, using wildcards * and

***** **matches any string or path as usual**

**matches paths only**

anim* - **would match anim, animal, anim/234, and animation:cell**

anim# - **would match anim and anim/234** **animal** - **would match animal only**

Querying for numbers:

*** has speed# = (> 115 kph)** - **would match cheetah has maximum speed = 72 mph, and arrow has speed:launch = 100 +/- 20 kph**

(> 100 kph && < 90 mph) is also a valid value expression. ***NOTE: numbers are converted to SI units internally for comparison***

Querying for strings or phrase paths:

Wildcards to not work on values that are strings or paths, like **= "How? Nobody knows."** **or** **= alert:high** These can be matched verbatim, or using a regular expression: *** has * = REGEXP("lert")** **or** **= REGEXP("obody")**

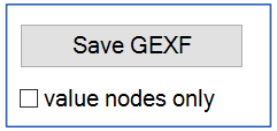
Tags: In queries, tags are paths preceded by the @ sign, such as **@status:deprecated** or **@status#** and can be grouped. **@[status#, !official]** will only return truenumbers having any **status** tag, and NOT having an **official** tag. A query can consist of only a tag or tag group, or tags can be specified after a **has** query.

Multi-line queries and variables: say you want truenumbers for all subject that have both a latitude and a longitude:

{ ?x has lattitude; ?x has longitude ; } matches them. **{ ?x has lattitude; ?x has longitude ; ?x has *; }** would additionally return all the other truenumbers for subjects that have both a lattitude and longitude.

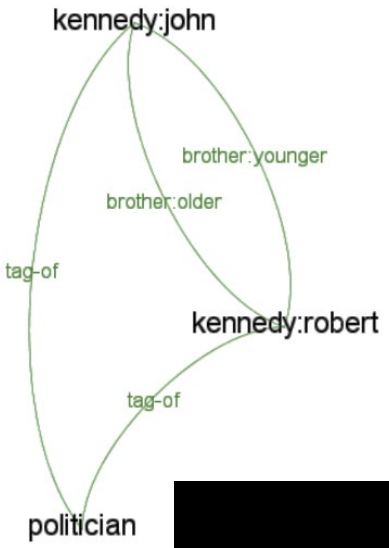
Generating GEXF graph files

GEXF (Graph Exchange XML Format) is used for describing graph structures. It specifies the nodes and edges of a graph as well as user-defined attributes such as node weights or edge directions. Intended as a general exchange format for graphs, it is used differently by different applications. Truenumbers generates GEXF files optimized for use with the online graph visualization and analysis tool **GraphXR®** (https://www.kineviz.com). They will also work with Gephi, and open source desktop tool (http://gephi.org).



Truenumber data is not graph data per-se but graphs are useful for visualizing aspects of a truenumber dataset. TN for Excel can currently generate two graphs from the set of truenumbers returned from a query. The default shows nodes for *subjects, values* and *tags*. Edges are labeled by the property connecting subject to value. Tags are connected to values with a "tag-of" relation. The "value nodes only" checkbox generates a set of nodes and no edges, one node for each truenumber, labeled by its value. Each node has attributes including its subject, property and tags. **GraphXR** has powerful tools for manipulating graph data allowing you to aggregate or generate new nodes and edges, so you can interactively create graphs. At left we have two simple truenumbers and the default graph generated for them as visualized in GEPHI 0.9.2. A slightly larger dataset is shown with "value nodes only" in the GraphXR tool. One node is selected, showing the properties it carries. We use the GraphXR extract and merge functions on the "subject_root" property, the root token of the truenumber's subject path, to generate a graph categorizing the numbers related to "Kennedy".

SUBJECT			PROPERTY		VALUE	UNITS	TAGS
kennedy	john	has	younger brother	=	kennedy:robert		politician
	robert	has	older brother	=	kennedy:john		politician



In most cases, using the default graph is a good starting point. Both Gephi and GraphXR allow edits like removing nodes, for creating exactly the diagram you need

.