

: *What is digital data?*

by Daniel J. Power

Editor, DSSResources.COM

Blog posts, tweets and high definition movies are all digitized. We live in a digital universe of bits and bytes. Digital data is 0s and 1s, discrete, discontinuous numbers or binary digits, readable by computing machines. Digital data, you may recall, is encoded using the binary numbering system. An 8-bit string of binary digits like 0101 1010 is an example of digital data. The binary number 0101 1010 can have meaning. For example, it is also the 8-bit ASCII Character Code for the uppercase English letter Z. Also, it is equivalent to the Decimal value 90 and the Hexidecimal value 5A. Please paste 0101 1010 into the Google search engine and confirm the coding and conversion. Most of us can't do the coding or conversion quickly without computer assistance. It is easy to forget that computers process data differently than we humans do.

In binary ASCII Code (cf., <http://www.ascii-code.com/>), my digital data name is 01000100 01100001 01101110 00100000 01010000 01101111 01110111 01100101 01110010 (aka Dan Power). Technically what you are reading is human readable data generated by a computer processor. Watching a memory dump of binary data on a computer screen provides the metaphor of a data stream. The bits stream across the screen very rapidly. Also, imagine packets of data streaming down the digital highway, the Internet.

Machines read, store and process both analog and binary data. Data can be converted between analog and digital. Video and sound recordings were initially stored as analog data. Current technology stores most data digitally. People can't easily read digital data, but computing machines covert digital data into human readable and understandable formats.

When facts, text strings, photos, or music are encoded using binary digits and hence digitized, the original data source is transformed into digital data. Computers, and people using them, directly create and generate digital data. According to Mearian (2012) based on an IDC study, "A majority of the information in the digital universe is entertainment or social media." The same IDC study found "by 2020, as much as 33% of all data will contain information that might be valuable if analyzed." Digital data is more than the bits and bytes a computer central processing unit reads, uses and can create.

So-called "raw" digital data, the bits and bytes, has no intrinsic meaning. Adding metadata is important. Bits must be processed into information, "bit" is a contraction of the phrase "binary digit". Processing adds meaning and context for both structured or unstructured data bits. In general, computer processing and especially database storage and processing is easier for well-structured data. Unstructured data like images, videos, email, documents and text, has no or a minimal

: *What is digital data?*

identifiable, repeated structure. Documents and text often have internally identifiable structures, but at best one would call such data items semi-structured or loosely structured. How do we store and retrieve the bits? A database management system (DBMS) is a software application for maintaining, storing, managing and retrieving digital data, especially structured data.

Both photos and moving visual images (video) are digital data. How are images translated into binary digits? Each pixel, a physical point in an image, is encoded in binary. Each pixel is assigned a tonal value (black, white, shades of gray or a color). We digitize the image and store, retrieve and process the encoded image (see answers.com). Software matches images by comparing the encoded pixels. That's how Facebook software can suggest the identity of a person in a photograph posted to the social network site. The software compares multiple instances of digital data looking for patterns and similarities.

We live in a digital world, but we do **NOT** need to think and count using binary digits. The binary system of counting can be so confusing to we humans. In binary, $10 + 10$ equals 4. So let's let the CPUs "think" in binary, while we humans continue to think in images and symbols, including numbers and text.

Digital data is the raw material computers capture, store and process. Bits become bytes and the digital data becomes a massive collection. How much data do we need and want? It seems the marketing folks, the police and the bankers and insurance salespeople all want more. They can't have too much data. Today we capture data about your driving habits, we monitor your financial transactions, and we can even monitor your health. More and better data is offered as the solution to every problem that humankind faces now and in the future. Digital data has value when used and companies want to capture and exploit that value. Digital data **is** valuable, but it can be misused.

So digital data is what computers read, write and process. We create and store enormous amounts of digital data. Each DVD holds 4.7 Gigabytes (GB) or the equivalent of 350,000 pages of 0s and 1s. Depending upon your assumptions, a Gigabyte is a quantity approximately equivalent to a pickup truck filled with printed digital data. We are swimming in a digital ocean of data. According to Gantz and Reinsel, our digital universe contained in 2011 "1.8 trillion gigabytes in 500 quadrillion 'files'" and data is more than doubling every two years. We live in a big digital data world and currently very little of that digital data is analyzed or secure.

Send comments to 01000100 01100001 01101110 00100000 01010000 01101111 01110111
01100101 01110010.

: *What is digital data?*

References

1GB = 1 Truckload at URL http://worldcadaccess.typepad.com/blog/2004/12/1gb_1_truckload.html

answers.com, "How are images and sound converted to binary?" at URL
http://wiki.answers.com/Q/How_are_images_and_sound_converted_to_binary

"Digital data," Businessdictionary.com at URL
<http://www.businessdictionary.com/definition/digital-data.html>

Data volume estimates and conversions at URL
<http://www.superiordocumentservices.com/resources/data-conversions/>

Gantz, J. and D. Reinsel, "Extracting Value from Chaos," EMC, June 2011
at <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>

Mearian, L., "By 2020, there will be 5,200 GB of data for every person on Earth," Computerworld,
December 12, 2012 at URL
[http://www.computerworld.com/s/article/9234563/By_2020_there_will_be_5_200_GB_of_data_for_e
very_person](http://www.computerworld.com/s/article/9234563/By_2020_there_will_be_5_200_GB_of_data_for_every_person)

Table of 8-bit ASCII Character Codes - Science Buddies at
http://www.sciencebuddies.org/science-fair-projects/project_ideas/CompSci_ASCII_Table.shtml

Wikipedia, "Digital data," at http://en.wikipedia.org/wiki/Digital_data

Author: Daniel Power
Last update: 2013-07-19 02:50