

The Trouble with Terms

Celia Paulsen
National Institute of Standards
and Technology

Editors:
Jeff Voas

Thousands of new words have been invented in the last decade to help us talk about technology. An analysis of the NIST computer security glossary database shows insights into how we invent and define these words, and the impact of those definitions.

Cookies are calorie-free; a mouse is not an animal; firewalls do not involve fire or walls, and the cloud has nothing to do with weather. Terms we use in the computer security / information security / cybersecurity world can create a significant amount of confusion.

By necessity, humans seem compelled to name things. Some of the first words a child learns are the names of things, beginning with “mom”, “dad”, or “apple” – things that can be seen and touched. Words for more intangible concepts like “hungry” or “ennui” come later. But it is these abstract terms that challenge the poet – finding the exact name for an emotion, a color, or an idea. Shakespeare found this process so challenging that he reportedly invented hundreds of words that are now commonplace in the English language (including addiction, lonely, and manager)¹.

With the advent of the Internet, a myriad of new words exploded into the public vocabulary. As related technologies, philosophies, and perspectives emerged, the technologist was faced with the poet’s dilemma: what name to give these *things*. New words were invented or old words repurposed. The Oxford English Dictionary shows that over 1815 words related to computing have been invented since 1950². In addition, the Merriam-Webster dictionary reported that many new terms are related to how information technology is managed, deployed, and organized³.

The National Institute of Standards and Technology (NIST) has created an online database (available at <https://csrc.nist.gov/glossary>) of glossaries published in NIST computer security-related publications and the Committee on National Security Systems Instruction (CNSSI) number 4009. An analysis of the glossary database provides some insights into how cybersecurity terms and definitions are created and evolve over time, not only at NIST, but broadly.

THE PURPOSE OF DEFINITION

In mathematics, definitions are created to shorten and clarify proofs. In language, definitions are used for similar reasons. Terms are created to encapsulate a complex emotion, thought, or concept into an easily recognized token. As such, words become a representative symbol. Without a

common understanding of the definition of a mathematical or linguistic symbol, communication becomes imprecise and, in some cases, impossible.

In early human history, measurements were taken by comparing an object to body parts such as the hand or a digit. Different sized people would measure an object differently. This caused problems in trade, engineering, and even farming. The Royal Egyptian cubit was the first known standard unit of length, reportedly responsible for the successful creation of the Great Pyramid⁴.

Reuven Cohen, cofounder of Cloud Camp, was quoted in a Technology Review article as saying, “The cloud is a metaphor for the Internet. That is why there is a raging debate. By virtue of being a metaphor, it’s open to different interpretations” and, “it’s worth money”.⁵ A term that is not clearly defined can mean many different things. This can have significant impacts in, for example, business when communicating the priorities of a project, in commerce when discussing the value of an object, and in law when discussing an arrangement.

The cubit provided Egyptians with a shared understanding of a measurement, allowing various groups of people to work together. Similarly, words can provide a standardized unit for measuring and communicating thoughts and ideas. Well thought-out and precise definitions enable a common, unambiguous understanding between differing parties, allowing them to measure and communicate concepts succinctly, clearly, and accurately.

THE MANY COLORS OF WORDS

The NIST glossary database contains over 3500 computer security-related terms from over 150 documents. An analysis of the terms found that many fall into one of two camps: abbreviations turned into words, and metaphors. In the abbreviation camp, *SIM* is an abbreviation of *Subscriber Identification Module*, *bit* is short for *binary digit*, and *captcha* is an acronym for *Completely Automated Public Turing test to tell Computers and Humans Apart*. Many seemingly made up or nonsensical computer terms were once abbreviations where the abbreviated forms became words in and of themselves, independent of their longer forms.

Thankfully, these terms often come pre-packaged with a definition, especially if the term is trademarked, such as with *wi-fi*. Richard Robinson in his seminal book *Definition*⁶ called these pre-defined definitions stipulated definitions. A stipulated definition is a “self-conscious setting up of the meaning-relation between some word and some object”. Robinson argues that stipulated definitions create concepts and are free from pre-connotations or associations based on how a word has been used before.

The value of stipulated definitions can be seen in terms that were created to represent a specific concept. Terms such as *hyperlink*, *kernel*, or *cryptography* have few unique definitions in the NIST glossary database. Stipulated definitions tend to be stable; there is generally common agreement as to what the terms mean, and they are less vulnerable to the cultural mood swings that tend to affect a definition.

Terms that are metaphors include *cloud computing*, *virus*, *bounce*, and *wizard*. These terms use words readers or listeners are familiar with and provide an instantaneous mental association that makes complex ideas seem accessible. Marketing departments love metaphorical terms for exactly that reason – they can allow the customer to interpret the term in a way that makes sense to them.

Unfortunately, the instinctual understanding that the metaphor provides does not automatically result in a usable, translatable, measurable definition. Technologists, standards groups, and legal groups are left to create a definition to match the intent of the words. These lexical or extracted definitions seek to explain how a word has been used and also provide value to the community, which can be challenging and inexact art.

LEGITIMATE DIFFERENCES

At times, existing definitions simply are not adequate. Existing definitions may be overly general, overly specific, or miss some key point, critical in the context of the publication where the

term is used. This may be due to advances in technology, expanding uses of technology, or a change in understanding.

The context of the use of a word is perhaps the most common reason for differences in definition. Understanding context is critical to understanding a term's meaning. A presentation about risk could be completely different depending on whether the person speaking is an insurance provider, a software engineer, or a project manager. The term *Risk* means different things to different people. Its use in NIST publications is tied to the context of the publication, for example referring to the risk of a breach, the risk of a failure, or the risk of a supply chain attack.

One of the most overt examples of this is in law. The legal definition for a term can be almost unrecognizable compared to the popular definition. For example, the terms *property*, *damage*, and *warrant* all have very specific legal definitions that do not always agree with the terms' usage in the common vernacular. Even more, the legal definition for a term can differ depending on whether the discussion is regarding patent law, family law, or finance law. The term *dependent* is one example. In patent law, *dependent* specifically refers to a claim with a reference in another claim; in family law, it is someone whose livelihood is sustained by another person; in finance, it can refer to a contract that is conditional upon a separate agreement. These definitions are similar, but tailored to the context.

As *privacy* becomes more important from a legal standpoint, the various legal definitions will likely not harmonize with the popular use of the term and will potentially also conflict with technical uses of the term in the computer world. This could cause confusion, especially for international organizations required to follow numerous regulations that may not standardize in something as fundamental as the definition of the word *privacy*.

This variability is not a fault or failure of the authors of the regulations, but rather a symptom of language itself. Words can have multiple definitions, either depending on context or as they change over time. *Hatch* in English can refer to a hatching egg, or the hatch of an airplane. Many years ago, the word *nice* meant *silly* or *foolish*, whereas now it means *agreeable*. Changes to definitions occur even in the most structured scientific fields. Although the term *planet* has been used for centuries, which celestial objects can be classified as a planet is an ongoing debate⁷.

THE LITERALIST

One of the more common reasons the meanings of terms change is due to misunderstandings between technologists and the lay person. When a word becomes very popular, marketers, media, academics, and others find they must use the term in order to garner attention, regardless of whether their use of the term is strictly accurate. Or, instead of remembering a new term for a new but similar technology, the populace may continue to use the old term, effectively changing its definition to include the new technology. When either of these scenarios happen, arguments between purists and populists can erupt, causing disagreement as to what a *blockchain*, *mouse*, *hacker*, or *emoticon* is.

One strategy to prevent such arguments is relating lay or common terms or phrases to more formal or scientific terms. For example, the phrase *superbug* is occasionally used by media to refer to *antimicrobial-resistant bacteria*, but the scientific community is generally strict in using the more scientific term. When discussing animals or plants, scientists regularly use the Latin (binomial nomenclature) name, but also mention the common name as a side note (e.g. "crataegus monogyna or common hawthorn"). This allows for the specificity that stipulated definitions provide, while connecting to the common lexicon.

At times, the literal or scientific definition of a term may differ so much from the popular version, that it can lead to significant misunderstandings. Parker Molloy wrote a comical but apropos article where she replaced the term *millennial* in headlines with the phrase *adults under 40*⁸. The results remove a lot of preconceptions and emotional biases associated with the term and afford greater understanding of what is actually being said. A potentially eye-opening experiment would be to perform this same exercise with headlines containing popular terms like *Internet of Things* or *Cloud Computing*.

SUMMARY

As humans fumble to communicate new, abstract, insubstantial and often seemingly ethereal concepts, they create new symbols (i.e. words) to represent difficult ideas. Ludwig Wittgenstein famously wrote, “The limits of my language means the limits of my world”⁹. Technology may also be limited by the language we use. Clear definitions of these symbols that are understood by all parties involved are necessary if one is to measure, manage, or hold accountable anyone or anything.

Developing a common, standardized understanding of computer-related terms could have broad impacts in computer engineering, contracting, and commerce as a whole. When a technologist, legal expert, and business executive all have the same understanding of a term, communication becomes quicker, easier, and more accurate. When developing and defining terms, the cybersecurity community should consider the usability and long-term measurability of the definition.

An analysis of a broader collection of terms and definitions would likely provide interesting commentary regarding which computer security-related terms have a general agreement as to their definitions, and which do not. NIST has published a draft report describing the NIST glossary in more detail (available at <https://csrc.nist.gov/publications/detail/nistir/7298/rev-3/draft>). Comments on the report and database are requested and can be sent to secglossary@nist.gov.

REFERENCES

1. <http://shakespeare-online.com/biography/wordsinvented.html>
2. <http://www.oed.com/timelines?scope=SENSE&subjectClass=Technology&type=timelines>
3. <https://www.merriam-webster.com/words-at-play/new-words-in-the-dictionary-feb-2017>
4. <http://www.msc-conf.com/history-of-metrology/>
5. <https://www.technologyreview.com/s/425970/who-coined-cloud-computing/>
6. Richard Robinson, *Definition* (1962). Clarendon Press. Oxford.
7. <https://www.space.com/25986-planet-definition.html>
8. <https://www.upworthy.com/i-swapped-out-millennials-in-headlines-for-something-better-it-made-a-huge-difference>
9. Ludwig Wittgenstein. *Tractatus Logico-Philosophicus* (1922)