

Labels are a necessity in the organization of
knowledge, but they also constrain our
understanding

Word count: 1283 words

February 15, 2021

Labeling knowledge allows us for more efficient learning and cooperation. In the age of the Internet, knowing how to ask questions and how different areas of knowledge connect are the skills that unlock the unlimited potential of the shared knowledge we have easy access to today. Labels are used to organize the information people have in their brains as well as in the external storage ranging from personal notes to huge data centers. The abstract structures in which the information is stored and added to has its drawbacks, as it makes it very easy to use these abstractions without the deep understanding of them which is a double-edged sword because it allows for a very rapid progress in a practicing a new skill or in a project, but also makes us lazy and easily satisfied with the illusion of understanding rather than actual deep knowledge in a subject. The main goal of labeling information is to retrieve it. Organization of information is being done on many levels. One is personal, where we use things like mnemotechniques that allow us to remember things like vocabulary of foreign language or physical formulas. Other one is on global level, where we have sources on information like the internet and books. Both of these levels work in tandem, because it is not sufficient to take in information through visual or audio stimuli, the information has to be processed and somehow associated with other data to be used effectively and creatively.

Labeling information has a great analogue in psychology - schema theory. Schema theory explains that information in brain is stored in a net of associations where one link can cause adjacent ones to activate too. Labels put on information does not float separately either, as every piece of knowledge has its connections to other related topics. Great example are physical formulas, say 2^{nd} law of Newton stating that $a = \frac{F}{m}$. Each letter in formula refers to other quantity, F to force, m to mass and a to acceleration. 2^{nd} law is the label for the idea that acceleration is proportional to the force applied to a body and inversely proportional to the mass of a body. This gives a compact form to communicate some way of solving a problem, say the time it takes a led ball to fall off of a roof of a house. The direct association one might have is that the ball is going to be falling with accelerated motion, so the scheme containing 2^{nd} law of Newton, which points to the formula, which contains acceleration, which in its description ($\frac{m}{s^2}$) contains time, which needs to be derived. Following similar steps, provided one has accurate map and labels, one can solve the problem not having the direct, trained solution to the problem, as it is quite inefficient to derive the same information over and over again. It is much easier to put all of the facts in one black box labeled “free falling objects” and then the solution is short-circuited to expression $h = v_0t + h_0 + \frac{1}{2}gt^2$, which can be the starting point. Such approach is efficient, as it is not always possible to understand the whole concept from the nuclear part to the desirable concept. For example calculus on high school level is presented more like a set of tools rather than a whole bottom-up process that can be understood every step of the way. This allows for “getting the job done”, but it only allows for solving explicitly stated problems and not problems that require going the way explained above and observing some discrete relations between concepts on much lower level of

abstraction. Such approach allows for optimization of already existing systems, but also for innovating, which, in terms of labels, could be stated as creating new labels and fitting them into the existing map of concepts.

One of the most obvious and most powerful uses of labeling information is the Internet. Search engines roughly simulate schemas explained above. When searching for a solution of a problem, the most common approach is to literally ask a question. What we do asking a question is feeding the search engine keywords, and, if we search for a very common question, like “What is the mass of the Earth?”, we are most likely to get a direct answer from our search engine, as it is such a ubiquitous expression, that is mapped directly to $5.972 \cdot 10^{24} kg$. There is a threat to such approach, as advanced search engines mutate the map of concepts to suit us best. This sounds great, but it puts us in a bubble of self reassurance, as we are fed the information we are most likely to agree with, and in turn consume more articles, videos, etc. This is especially prevalent in political queries, as this is one of the most polarizing fields. If we are, however, to search for more specific and less popular things, we can better observe how search engine works, as we get results based on the keywords in our question. This puts the person searching in a similar position as the student learning calculus, as it gives a starting point to get something done, rather than the understanding to solve the problem on one’s own. This gives an idea of how learning from the Internet can stifle understanding and give weak foundations for problem solving, as it is very easy to learn to do something without making connections to other concepts. Conversely, learning from the Internet can be done with the emphasis on understanding with the web being an external media for nuclear parts of concepts fitting on the ends of links.

Uses of labeling can be also misleading in it of itself. Labeling enforces some structure, which not always is the best for observing the applications and relations between different domains. Such example is mathematics, where an expression “I will never, ever use it” coming from students is a cliché. This is because maths is as a tool without anything practical to work on, which does not form any connections to other fields and when faced with a problem from other domain, it could make it difficult to make the connection between the motion of pendulum to sine wave. All of this is the other side of the coin of focusing on focusing on nuclear parts, as without them being enclosed in boxes, it is more difficult to understand applications in different fields. This could be defined as more entrepreneurial approach where what matters is finding an easier solution to any problem. This, of course, does not apply to business only, as the mere relationship between pure and applied mathematics. An example could be Fourier series, which is a way to use sums of sine and cosine waves to model any periodic function. This came from the need to model the heat transfer function, which has a “step” in it. All in all what Fourier did was ask a question of how to solve a problem with existing set of mathematical tools, and understanding concepts from very low level of abstraction, allowed for deriving a method that is used all over physics and computer science 200 years later.

Labeling of information is crucial to learning and teaching. However, there is a delicate balance to understanding links between big sets of concepts, which could be AOK's, and nuclear parts of these sets, like 2nd law of Newton in physics. Labels themselves do little to enforce understanding, they merely give directions. Understanding is the skill of navigating through these concepts and having as many relevant connections to other fields and facts as possible on as many levels as possible.

Sources

- Fourier Series: https://en.wikipedia.org/wiki/Fourier_series
- Fourier: https://en.wikipedia.org/wiki/Joseph_Fourier