

Content Optimization

By Hans Fenstermacher

Why Technology Can't Do the Job

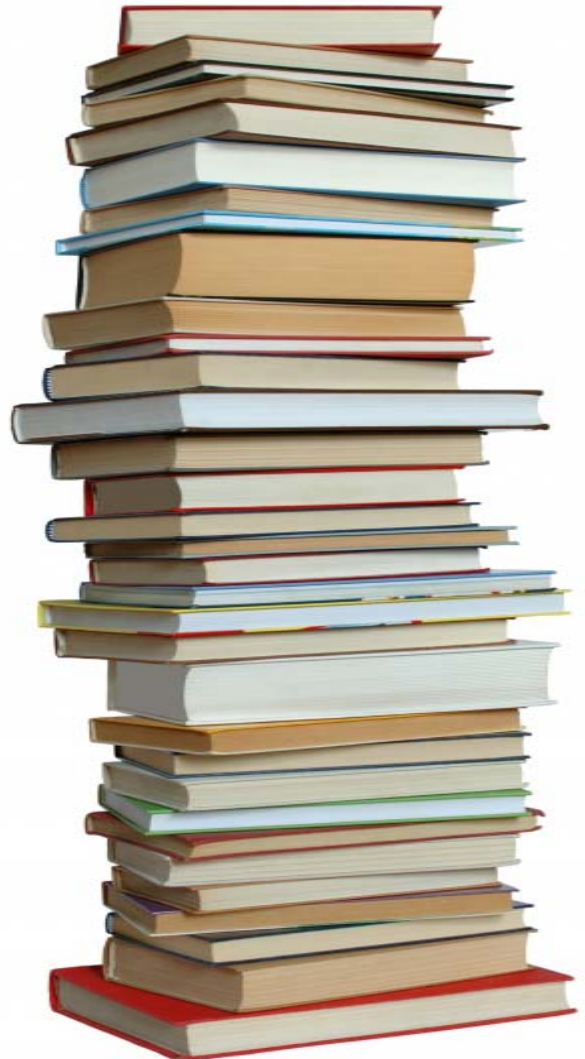
The technical authoring and localization industry seems to focus on technology as the primary solution to most problems. In a way, who can blame us? After all, the high-tech sector generally has been the leader in innovation. Our industry tools have significantly boosted productivity and efficiency. Now that content optimization looms large as the next big problem to tackle, many of us are naturally depending on technology to solve this problem, too.

Unfortunately, we're headed for frustration. Technology cannot lead the content optimization mission, because the biggest problems in content aren't technical. When it comes to content, technology has fundamental shortcomings. The sooner we realize this truth, the better we will be able to solve the challenges before us in real ways.

Let's look at some of those shortcomings.

Technology can't resolve ambiguity.

Natural language, the primary medium for all content, is full of ambiguity. Technical authors do everything they can to avoid it, but they don't always succeed. Take this sentence: "The label is on either side of the drum." This could mean that there are two labels, one on each side



of the drum, or it could mean there is only one label, which may be on one side of the drum or the other. Without more information, either interpretation is correct, but both cannot be. (Side note: This example comes from real documentation. When asked which interpretation was correct, the author responded, "I don't know!").

There is no technology or tool that can resolve this kind of ambiguity, because human beings can't resolve it either. The sentence is grammatically correct—it may even be terminologically correct—but it isn't unequivocal. For translators, this means they must choose one of the two interpretations. Given this 50-50 probability, it's likely that half the translations will be wrong.

Technology doesn't understand anything.

Another thing technology can't do is actually think. Tools often give the impression that they are thinking, but they are merely comparing things (occurrences, odds, rules, etc.). Consider the following sentence: "The main purpose of the payroll run is to calculate employee pay correctly at regular intervals." From this sentence, a content-analysis tool might "understand" that payroll run is a process, that this process involves an action (calculate employee pay), and that this action occurs with some frequency (at regular intervals). Fine. But what the tool doesn't know is that this sentence is completely uninformative to actual human beings; anyone who's ever held a regular-paying job knows what the purpose of a payroll run is.

By not understanding what it is processing, technology fails to tackle one of the most fundamental problems in content optimization: usability. Technology can make content consistent, more uniform, more structured, and (technically, perhaps) error-free, but it cannot make content informative. Only a human editor knows what this sentence really needs—to cease to exist.

Technology doesn't care about usability.

Technology doesn't even know what usability is. Tools that manage content structure and workflow are designed to follow rules (ironically, rules that must be created by human beings), the outcomes of which are binary—yes/no, allowed/prohibited, accepted/rejected, 1/0. But usability isn't binary; it's full of complex issues like context, exceptions, interpretations, and so on. The creators of content (humans) are not binary, and the recipients of content (humans) are not binary. Why would we expect the binary technology in between to ensure usability?

Usability requires understanding who the recipients of information are and what they

need. Consider again the payroll run sentence: all the binary rules may be followed precisely for formulating this information, but only the creators and recipients know that the information is useless (usability = 0). No technology is actually intelligent enough to delete this sentence.

Technology can't restructure content.

Content structure is a key to its usability. One of the critical tasks technical authors perform is analyzing and organizing information to make it usable. Technology simply can't do that. Take this example:

"To embed a graphic in a text document, you must open the object in its native application, remove all font information (for example, converting fonts to outlines), and save the file in JPEG format. Then in the text

document, place your cursor where you want the graphic, and use Import to insert the JPEG file you just created.”

This highly unreadable (dare I say off-putting?) text block becomes usable with good structure:

To embed a graphic in a text document:

1. Open the object in its native application.
2. Remove all font information (for example, convert fonts to outlines).
3. Save the file in JPEG format.
4. Open the text document.
5. Place your cursor where you want the graphic.
6. Use Import to insert the JPEG file you just created.

Technical authors apply these techniques almost instinctively, but tools don't have a clue.

Technology can't self-correct.

Tools are very precise; that is their strength. Tools also rely on precision implicitly; that is their weakness. All content processed by tools originates from a human being, and human beings are by nature imprecise. This imprecision frequently exceeds technology's tolerances, but it's far less of a problem for other humans. How many animals of each type did Moses take on the ark with him? The correct answer is none, because it was Noah who built the ark, not Moses. Many of you, though, undoubtedly self-corrected that error without even knowing it.

Authors of technical content are trained to be as precise as possible, but nobody's perfect. Minor errors are quite commonplace, as these examples show (corrections provided in parentheses):

- Then lick the Date/Time Stamp button. (click)
- The air damn directs the airflow. (air dam)
- You must not turn off the printer. (now)

Relying on spell-checkers (which also demand complete precision) cannot save the above or countless other examples of misinformation. Incidentally, after localization, this information can no longer be self-corrected by readers, because the cause of the imprecision is literally lost in translation.

Technology can't reduce volume.

Strictly speaking, tools can limit volume by imposing constraints. Proponents of text analysis and processing tools say that volume limits demonstrate the value of the technology (you know, maximum 7 words per sentence, that kind of thing). But artificial limits do little to optimize content. The philosopher Bertrand Russell famously showed that a few words can be just as uninformative as many:

Colorless, green ideas sleep furiously.

Real volume reduction requires a deep understanding of content: How can text be restructured, combined, and rearranged to make it shorter? What information is unnecessary because it is superfluous or redundant (remember: redundancy is a hallmark of technology)? Only human beings can answer these questions. If fact, users (also human beings) urge us to do so every day by refusing to read the documentation we produce.

Technology isn't natural.

Last, but not least, let us not underestimate human nature. Out of the last 2 million years of our evolution, we have spent—statistically speaking—approximately 0% of that time in the digital age. So, every technology is designed to mimic the analog world, because if something doesn't look and feel "natural," we won't use it. Writing is a deeply analog task, and humans fiercely resist writing techniques that aren't natural. That's why controlled language hasn't gained much in popularity among technical authors in over 30 years; it just runs too deeply against every natural instinct writers have. For technology to truly succeed, it must feel natural, so we will want to use it.

Has your car ever driven you to work? Has your washer ever done the laundry for you? Have hardware and software ever gotten together and written a user manual while you weren't looking? Of course not. They need you to do those things.

Technology is only one of the tools for content optimization. Remember Abraham Maslow's dictum: If all you have is a hammer, you tend to see every problem as a nail.

The problem of content optimization is complex, and the problem requires complex, not pie-in-the-sky, solutions. Content management and single-sourcing proponents often hold the view that technology will solve the content optimization problem. They seem to think that optimizing the workflow around content is the same thing as optimizing the content itself. Anyone who has ever read a user manual, however, knows that just isn't true.

About the Author:

Hans is president and founder of ArchiText Inc., a globalization and localization services company, and founding Chairman of the Board of GALA. At ArchiText he has pioneered ABREVE Connect, a globalization services and training program that helps companies implement a global content strategy and save localization costs.

