
How to Read Copious Texts Effectively

If you are an executive in a high flying team, whose task it is today to go through 10,000 survey responses to understand major trends, or to read 100 analyst reports to figure out if anything interesting has happened, we sympathize. For more than two years now we have been training Coseer's computers to help you out. Recently we paused, and realized that a cognitive computer can also teach us a lot about how to read copious texts. In this blog we have shared what might be the first lessons humans learn from computers.

1. Use the What-and-Why framework

After trying out multiple semantic models we have settled on the point of view that synthesizing any text is basically equivalent of asking three questions. We call this What-and-Why framework.

1. What are the most important topics?
2. Why are these topics important?
3. Does the response to Why for a given What need recursive What-and-Why processing to better understand it?

#1 identifies the major topics talked about in the text; #2 gives a topic more color and talks about the reasons building up to its importance; and #3 takes the responses from #2 and further decomposes in subtopics and their reasons to be important, essentially creating ontologies and structures where necessary. For a computer this framework quickly converges around critical things, ignoring the rest. There is no reason for that not to happen with humans - keep reading!

2. Always know Whats before looking for Whys

Coseer looks at information incrementally i.e. it analyzes each sentence one at a time and updates scores of variables pertaining to the sentence itself, the topics the sentence talks about and the relationships it implies. Even for a computer, which can easily manage large arrays, the process becomes unscalable beyond a point. After scrambling for different techniques it turned out that Coseer can simply ignore sentences that are in no way connected to topics of current interest. For example, if a user is trying to know more about Apple Inc, there is no point processing a sentence about Antarctica unless the two are somehow explicitly connected.

The human mind can effectively store and process only a few things at a time. You should ignore more things than Coseer to be effective. If you already know what you are looking for, you can quickly scan through the text, one paragraph or one sentence at a time and stop to engage only when you see the keywords of your interest. Most of texts today are digital, and come with a search button. Just hit F3 as many times as necessary.

Essentially, there is no point trying to find Whys when you don't know the Whats.

3. Generate Whats if there are none

In most cases you already have a good sense about the Whats - call it domain expertise, the alert in your inbox, or just a good hunch. In some cases you may be genuinely curious or may not want to bias yourselves. Even then, you should invest in finding the Whats before you go into Whys to consume the text effectively.

A well structured text offers clues to the Whats in its titles, subtitles and use of Proper Nouns. These signals are available to computers and humans alike. The table of contents is a great starting point.

For non-structured text, e.g. survey responses or reviews, Coseer has a slight advantage. It can parse the text into noun phrases, verb phrases and such, weigh every phrase and quickly converge onto the most important Whats, at least for the first iteration, all in milliseconds. You can still use a computer to help you without running advanced artificial intelligence algorithms. Try feeding your entire corpus into simple word count programs and then use your judgment and domain knowledge to form your hypotheses.

4. Keep change constant

Unlike Coseer's computers, you possess judgment. For example, Coseer has to jump through many hoops to decide that the first sentence is far more important for the topic "John" than the second one.

- John is an ace student.
- The class has three students whose names start with J - Jessica, John and Judy.

You just know. So, unlike Coseer you can easily figure out that if all instances of John are in sentences like the second one, it's time to drop John as a What and focus on something else. We have found that such mid-course corrections can increase efficacy of the exercise by a lot. In fact we are spending a lot of time training Coseer to make such decisions.

5. Abandon the fear of missing out

After the first pass of the What-and-Why reading method, you should have captured the most important themes in the text and the most important aspects for these themes. At this point, Coseer's computers do a second pass, and a third, may be a fourth, as many times as needed. Armed with your first iteration hypotheses you can also make very quick scans again. We posit that the time taken for both passes will be less than a traditional single pass. The information captured will be significantly better.

In fact, we learnt that we can go a step ahead. We learnt that in most cases the second pass is completely unnecessary. It's true that a second pass will add to the notes you have been taking, but this information is hardly ever useful. Ergo, if you can conquer your fear of missing out, you can really reduce the time taken to read copious amounts of text. Granted, it's an emotional matter.

Happy reading. Do let us know if this worked for you. Today maybe you can catch a matinee.

What is Next-Generation Enterprise Search?

Coseer's search solutions are transforming industries from healthcare to finance. Our point-and-shoot AI trains finds answers and insights with 95%+ accuracy within 4-12 weeks - all of this in 100% security. The reason? We founded Coseer on the principle that computers should take care of the boring stuff so that humans can focus on creativity and judgment. To that end, we've built enterprise search solutions to complete complex workflows just as humans would in a fraction of the time. Fortune 500 leaders are using Coseer to speed up and automate their most complex work.

We follow a tactical approach to enterprise search:

- We deliver 95-98% accurate solutions within 4-12 weeks.
- Our solutions deploy entirely behind your own firewall for 100% security, and every decision point is logged for full transparency.
- You add the finishing touches, but our point-and-shoot AI practically trains itself. No more huge training data sets or time wasted annotating and tagging.

Visit our [website](#) for in-depth case studies, ROI breakdowns per industry, and other insight.