DOI:10.1145/3483539

Vinton G. Cerf

The Future of Text Redux

I HAVE WRITTEN ABOUT text in its digital form in the past and I would like to revisit this topic once more. J.C.R. Licklider and Douglas Engelbart were two giants who saw non-numeric possibilities in networked computing well ahead of many others. Vannevar Bush and Ted Nelson are two others who resonated with the idea of machines that assisted in the production and discovery of information. Sir Tim Berners-Lee amplified some of these ideas with the invention of the World Wide Web. Following along this path is Frode Hegland, a protégé of the late Douglas Engelbart, who has developed new tools for the production of and interaction with text. Engelbart's oNLine System (NLS)^a was a tour-de-force example of disciplined use of structure to guide the production and consumption of computer-based text. One could view content at various depths (for example, first line of each paragraph, first paragraph only, subsets of segments of text found in classic structured outlines of document) and one could accomplish major restructuring of documents with ease because the NLS understood the structure and provided ways to reference portions of documents to facilitate restructuring.

Hegland has developed three remarkable tools for text generation, viewing, and referencing that inherit some of the philosophical aspects of NLS and enrich them with more fluid ways of organizing, viewing, and generating text. His three contributions are AUTHOR, READER, and VISUAL-META.^b These three tools illustrate

the power of applying computing to text, creating lenses through which to create, consume, and reference content. Hegland's focus is less on the appearance of text, over which most text editors persevere, than on its structure and relationships among various parts of a document.

Hegland's AUTHOR program allows the producer of text to visualize and manipulate it in other than linear ways. The approach supports concept-focused writing, allowing the user to define as they write, and to then see the defined text in a Concept Map, while also exporting all defined text as an interactive Glossary.

The major contribution presented here is VISUAL-META. This is an approach for adding metadata to PDF documents, in a form that is equally readable to human and machine. Such data can then be interpreted and used to create properly formatted citations and afford more elaborate manipulations such as interactive graphs and charts. Hegland's insight is to add this information at the end of a PDF document and give it equal stature as the text of the document itself. In a sense, such a document "knows" itself and uses that knowledge to assist a reader. It is important that it enable both the human reader and software reader. VISUAL-META is entirely open for anyone to employ, with full specification being available on how VISUAL-META can contain citing information, structural/heading, glossary, endnotes, references and more, at the project website: <http://visual-meta.info>

Visual-Meta enables PDF readers, such as Hegland's READER, to augment the reader's ability to consume text in a variety of ways, such as the ability to copy text from the PDF and paste it in a Visual-Meta enabled

word processor, such as AUTHOR, where it will appear as a full citation, reducing the chance of errors. It also allows for novel views which basic PDF does not provide, both for a single document and for large volumes of documents.

These works are instances of what I think should be called *computational text*, by which I mean, text that lends itself to augmentation through computational tools. Engelbart referred to his research laboratory as the Augmentation Research Center^c to emphasize the capacity of computers to augment human capabilities and the concept seems applicable to augmenting the utility of text. One begins to visualize galaxies of content in a mixed-media universe and tools for production, discovery, and consumption. "But isn't that just the World Wide Web?" you might ask. Well, yes and no. The Web is indeed a remarkable and linked universe of wide-ranging content. Search engines and browsers aid in our ability to find and consume, render, and interact with that content. Hegland's tools add an exploitable self-contained self-awareness within some of the objects in this universe and increase their enduring referenceability.

It is exciting to learn the ACM Digital Library is exploring aspects of the VISUAL-META concept for incorporation into its operation. But then, one would expect ACM to look to cutting-edge ideas for the benefit of its members. □

^c https://en.wikipedia.org/wiki/Augmentation_Research_Center

Vinton G. Cerf is vice president and Chief Internet Evangelist at Google. He served as ACM president from 2012–2014.

Copyright held by author.

^a [https://en.wikipedia.org/wiki/NLS_\(computer_system\)](https://en.wikipedia.org/wiki/NLS_(computer_system))

^b <https://www.augmentedtext.info/>