# Fatal Flaws in ChatAI as a Content Generator

**Hal Berghel**, University of Nevada, Las Vegas

*Unfortunately, the current fascination with AIChat focuses on content generation, which may be its weakest suit.*

**T**here is a lot to like about AIChat. For one thing, it satisfies human curiosities, especially relating to some of Alan Turing's observations about the potential for machine intelligence. In general, and despite religious proclamations to the contrary, curiosity is a great motivator.[1] As Nietzsche put it, "What does not kill me makes me stronger."[2] What is more, AIChat is fun! But beyond that, it may it may ultimately prove valuable in repurposing codebases, as a component of database query support languages, extending the capabilities of scene analyzers and pattern recognizers, getting the sense of the contents of massive databases, etc. Large language data models have already shown themselves as useful utilities for suggesting linguistic alternatives, additional contexts, viewpoints, facts, etc., in text processing. It would also seem to have great potential in natural language translation. It is unfortunate, however, that the

current fascination with AIChat is automated content generation, for this is one application in which AIChat has the least to offer. As we suggested elsewhere,[9] although AIChat generally, and ChatAI in particular, may be said to pass a limited Turing test given some important caveats, this may say more about the limits of Turing tests than it does about the strength of ChatAI. I look at AIChat output as a major step below hastily authored polemics, self-published books, and gratuitous, vanity websites. In fact, to carry this analogy further, there is a sense in which AIChat is vanity publishing at its worst, where any vainglory lingers in the anonymity of the cloud. In terms of content, AIChat is better thought of as bloviation generation. As such it is likely to contribute little to edification and scholarship, although it will have profound effects on social media and partisan politics.

So what is wrong with AIChat? Why will it be of limited value to scholarship? In short, AIChat has some serious design flaws. For one, AIChat fails to advance interactive technology much beyond the more pedestrian goals of content generation. It is masterfully implemented, to be sure, but impoverished when viewed in the light of the visions of futurists like Vannevar Bush, Doug Engelbart, and Ted Nelson. Such visions (e.g., Bush's memex, Engelbart's oN-Line System, and Nelson's Xanadu) were focused

EDITOR **HAL BERGHEL**
University of Nevada, Las Vegas; hlb@computer.org

on achieving worthy, substantial intellectual goals and not shallow, uninformed, pedestrian activities. I expect that Ted Nelson (the last of these visionaries standing at this writing) is nonplussed by AIChat's feeble attempt to advance our collective intellect. In fact, the deficiencies of AIChat in fact parallel those of the World Wide Web. This is no coincidence.

## THE WEB

The World Wide Web was spawned by the work of Tim Berners-Lee at the international particle physics laboratory, CERN, in Geneva, Switzerland, and significantly advanced by the Mosaic browser software developed by Marc Andreessen and Eric Bina at the National Center for Supercomputing Applications at the University of Illinois in 1992.[10] The defining characteristics of the web were twofold:

1. the deployment of the new platform-independent, application-layer, communication protocol called HTTP that supported extensible, distributed, collaborative hypermedia
2. the integration of a variation of the SGML tagged image format language standard called HTML, which integrated hypertext into documents/web pages.

These two contributions laid the groundwork for the open, interactive Internet information system that we call the World Wide Web. As such, the web may be thought of as a paradigm shift away from earlier, more insular, application-layer protocols.

The idea of using a tagged image format for markup languages preceded the development of HTML by several decades. Both SGML and HTML descended from IBM's GML language standard defined in the late 1960s.

Certainly, the addition of hyperlinks into HTML was noteworthy. Hyperlinks enabled network resources to be accessed easily through uniform resource locators (URLs). Thus, the use of hyperlinks in HTML took the GML concept one step closer to the vision

---

> In fact, to carry this analogy further, there is a sense in which AIChat is vanity publishing at its worst, where any vainglory lingers in the anonymity of the cloud.

---

of linking together the world's information resources. Browsers added considerably to the value of the web by supporting robust media rendering and providing convenient techniques to access web resources via a mouse click on sensitized and clearly identified links in rendered media.

For these reasons, the web became a primary enabling technology on the Internet, quickly taking its rightful place along with earlier "killer apps" like Telnet, FTP, and e-mail.

## THE NATURE OF THE PARADIGM SHIFT

So how far did the web go in realizing the visions of Bush, Engelbart, and Nelson? In short—and this is the most important point—it did not go nearly as far as the hype and hyperbole would suggest. The advantages of the web over earlier application-layer protocols included 1) open architecture, 2) platform and media independence, 3) native web-compliant browser support for convenient Internet navigation and robust media rendering, and 4) the use of unidirectional hyperlinks. These advantages alone are not the stuff of which cybervisions are made.

While it certainly must be admitted that the web constituted a considerable practical advance in Internet

information access and delivery systems, it was conceptually primitive when compared with the visions of Bush, Engelbart, and Nelson. For one, the implementation of the hyperlinks was crude. As implemented on the web, hyperlinks involved drawing file-oriented chunks of media to the browser for independent rendering as distinct entities. There was (and remains for the most part) no correlative capacity for integration and annotation. So where you might have started with one web page, after clicking on a sensitizes hyperlink, and through the magic of HTTP and HTML, now you have two web pages. This linear accumulation of independent resources is at most a minor technological leap toward information absorption. While the web was comparatively a major advance in network media sharing over FTP and Telnet, it falls far short of the visions that Bush et al. had in mind.

I will illustrate the deficiency by way of this quote from Vannevar Bush:

> "Our ineptitude in getting at the record is largely caused by the artificiality of systems of indexing.... The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain.... Selection by association, rather than indexing, may yet be mechanized.

A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.... [memex provides] associative indexing, the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another....

Before him are the two items to be joined, projected onto adjacent viewing positions. At the bottom of each there are a number of blank code spaces, and a pointer is set to indicate one of these on each item. The user taps a single key, and the items are permanently joined

Thereafter, at any time, when one of these items is in view, the other can be instantly recalled ... Moreover, when numerous items have been thus joined together to form a trail, they can be reviewed in turn.... It is exactly as though the physical items had been gathered together from widely separated sources and bound together to form a new book. It is more than this, for any item can be joined into numerous trails...Next, in a history, he finds another pertinent item, and ties the two together. Thus he goes, building a trail of many items. Occasionally he inserts a comment of his own, either linking it into the main trail or joining it by a side trail to a particular item."[3]

It is important to remember that Bush wrote this in 1945, before commercial computers, before the Internet, before text and word processing programs, and before the formal study of information retrieval (IR). But it is fairly easy to see that Bush is committed to the principle of bringing together concepts and ideas, which goes far beyond linking static documents. The trails of which he speaks are integrated, annotated threads of fragments from many sources, carefully woven into a new creative artifact. Note that this does not involve the linear transversal of a sequence of hyperlinked documents currently provided by the web, which falls quite short of his vision. One does not create an original manuscript by assembling a collection of old readings.

## EXTENDING BUSH'S VISION

Flash forward twenty years, and Ted Nelson introduces the world to hypertext, not the hobbled form that we know from the World Wide Web but a more full-bodied form implied by Bush. Nelson's planned version of hypertext directly supports content transclusion and copyright management still absent from the web.

"If a writer is really to be helped by an automated system, it ought to do more than retype and transpose: it should stand by him during the early periods of muddled confusion, when his ideas are scraps, fragments, phrases, and contradictory overall designs. And it must help him through to the final draft with every feasible mechanical aid—making the fragments easy to find, and making easier the tentative sequencing and juxtaposing and comparing."[5]

He expands his idea with an overview of the overall design that a worthy information sharing system, which Nelson named evolutionary file system (ELF), should have:

1. There must be support of hyperlinks to network file resources.
2. There must be informal, not rigid, formal file relations—it must store media in any form and arrangement desired.
3. All files must be able to hold associated commentaries and explanations.
4. The files and file elements must be modifiable at will, and such modifications must be "dynamic" in that the change of one element will automatically change other associated elements.
5. The files must be retained in a version-controlled environment.
6. All element links must be bidirectional.
7. The organization of data elements must support nonlinearity.

We note that only requirement 1 is supported by the web, and even then in only limited form, for 1) web hyperlinks are unidirectional and 2) most links remain URLs and not URNs (resource names) or URIs (resource identifiers). One sees a clearer image of Nelson's ambitions in his definition of hypertext as a "body of written or pictorial material interconnected in such a complex way that it could not conveniently be presented or represented on paper. It may contain summaries, or maps of its contents and their interrelations; it may contain annotations, additions and footnotes from scholars who have examined it [and leading to] is a bundle of relationships subject to all kinds of twists, inversions, involutions and arrangement: these changes are frequent but unpredictable." Obviously, Nelson has in mind a data structure that supports collaboration, specifically including transdocument interconnectivity which he elsewhere calls transclusion.[6] His intentions, which are consistent with Bush's original vision, are clearly seen in Figs. 2–4 in his articles.

Nelson, like Bush, is short on technical details in his publications. However, if computer scientists take the time to seriously study these contributions, they will see that the scope of their collective visions (along with

Engelbart's[7]) circumscribes a much more expansive view of IR that spans networking, data processing, text and word processing, and multimedia. The overarching idea is that human-centered IR must be driven by the intellectual requirements of the users. In modern terms, we might label this the search for a grand unification of all cybermedia under a robust set of content-oriented, resource-level protocols. In any case, such requirements must efficiently and effectively advance human creativity in all cognitive areas, including science, technology, art, humanities, etc. While the articles may seem obtuse by modern standards, the terminology archaic (at least in terms of computing history), and the descriptions and definitions awkward, the underlying concepts remain innovative, radical, and consequential to the present day, well over a half-century after they were advanced.

## THE VISION, EXTENDED

Doug Engelbart summarizes the goal of the collective Bush–Engelbart–Nelson vision nicely in his companion to Nelson's 1995 update.[4] What is needed, he offers, is a tool that focuses on a universal knowledge base rather than function-oriented paradigms like word processors, file transfer protocols, and even the present manifestation of hypertext! The differences with existing systems become clear in his discussion of relevant data structures:

"Every knowledge object—from the largest documents, to aggregate branches, down to content units such as characters—has an unambiguous address, understandable and readable by a user, and referenceable anywhere in the hyperdocument system....A structured, mixed-object hyperdocument may be displayed with a flexible choice of viewing options: selective level clipping, filtering on content, truncation or other transformation of object content, new sequences or groupings of objects including those residing in other documents, etc. Links may specify views so traversal retrieves the destination object with a prespecified presentation view (e.g., display as a high-level outline or display only a particular statement). View specification becomes a natural and constantly employed part of a user's vocabulary."

It is clear that both Engelbart and Nelson advance the idea that the current web, browsers, and media-manipulating technologies should be seen as merely a first step toward a much more ambitious cybermedia-oriented end state. Two mission-critical features necessary in realizing this potential are Nelson's transclusion and transcopyright. Transclusion is particularly interesting as it involves "reuse with original context available, through embedded shared instancing" that spans any arbitrary number of sources and at an arbitrary level of element size within artifacts. The operational principle is the integration of contexts in the presentation and manipulation of media. The concept is easier to understand with one of Nelson's demonstrations.[7]

The many talks and demonstrations provided by Nelson and Engelbart collectively circumscribe the virtual corpus of their vision or, more accurately, the vision they shared with Vannevar Bush. The term "virtual" must be emphasized because neither was able to bring their vision to commercial fruition, which is a measure of both the complexity of the task and the limited immediate business appeal. Such was our experience with the enterprises of detecting gravitational waves and proving Fermat's last theorem. One must remember that the challenge of implementing a namespace for consistent and complete uniform resource names (URNs) proved so daunting that the entire enterprise was abandoned in 2005. The fabrics of transclusion and transcopyright infrastructures are of a distinctly greater complexity than even URNs. The evolution of the web after the usurpation of control by corporate interests is a history of picking the lowest-hanging and easiest to implement Internet fruit with the greatest commercial potential (compare the fates of cookies and shopping carts versus the do-not-track HTTP header). This is not to deny the enormous contributions of the Internet and the World Wide Web in support of knowledge workers. But both, and the web in particular, are only instruments to advance connectivity to global digital resources. They are passive with respect to advancing human intellect. They connect people with resources, plain and simple. Some Internet resources are noble and worthy, while others are unfit for human consumption. The web is agnostic in this regard.

## AUGMENTING HUMAN INTELLECT VERSUS AUTOMATING BLOVIATION

Bush, Nelson, and Engelbart were all about the use of technology to increase the efficiency and effectiveness of human intelligence, as opposed to the development of nonhuman intelligence. Of course, using digital technology to interconnect digital resources and build thought swarms and collaboratories on the back of nonlinear media traversal and independently of the logical structure of the forms of media, and so forth, were one part of this mix, but only a part. But there is much, much more involved with the process of augmenting human intelligence than networking computer systems, just as there is much more to text processing than the cut–copy–paste desktop metaphor and the rapid cursor movement technologies. Doug Engelbart, in particular, understood this point as he made contributions to both the vision and the mundane.[11,12] The working model of the vision must involve the creation of an environment that facilitates the creative enterprise, but the model environment is not the end goal. That is the major deficiency with the content generation aspect of AIChat, which involves

a retro dorsal attachment to the static, web/Internet model of resource interconnectivity rather than the vision of intellectual augmentation proposed by Bush, Nelson, and Engelbart.

To illustrate, if a serious commitment was made by AIChat to enhance human intellect, the absence of source identification and credit (as in Nelson's transcopyright) should have been an immediate deal breaker. The idea that human knowledge may be advanced by purloining anonymous content from undisclosed data repositories, with or without the use of large language neural networks, is preposterous. This is not "standing on the shoulders of giants" but more akin to wallowing in the muck and mire with lower life forms. This is precisely what is wrong with nonvetted recommender systems: without knowing the source of recommendations, there is no way to assign credibility and value to the recommendations. This is also the primary reason to challenge any lingering faith in the wisdom of crowds.[13]

And so it is with nonvetted automated content generation. What are the sources of this content? Absent identification, documentation, and context, such content is simply automated. anonymized blather. What could be wrong with building content generation on that?

The underlying philosophy of AIChat is inherently antischolarly and predicated on the principle that the value of content to an audience is determined by tribal serviceability. We observe that this is one legacy of partisan politics as we know it. It should surprise no one that AIChat will prove to be extremely popular in that realm.

Will AIChat (qua automated bloviation) become a displacing technology? If history is any measure, the opportunity to advance content-light, undocumented, and uncredited opinion to a premier position in mass communication is a temptation too great to be passed over by manipulators of the public, so the likelihood is that the answer will be affirmative. By its very nature, resulting communication will be divergent rather than convergent, and it is unlikely to produce any great innovations or unique capabilities.

On the other hand, although AIChat is arguably subcerebral knowledge work, it has obvious, residual potential regarding information management, specifically in terms of applications, interface support, IR, code repurposing, and the like. ▣

## REFERENCES

1. J. Agar, "2016 Wilkins–Bernal–Medawar lecture the curious history of curiosity-driven research," *Notes Rec., Roy. Soc. J. Hist. Sci.*, vol. 71, no. 4, pp. 409–429, Dec. 2017, doi: 10.1098/rsnr.2017.0034.

2. K. Stoner, "Science proves that what doesn't kill you makes you stronger," *Northwestern Now*, Oct. 2019. [Online]. Available: https://news.northwestern.edu/stories/2019/10/science-proves-that-what-doesnt-kill-you-makes-you-stronger/?fj=1

3. V. Bush, "As we may think," *Atlantic Monthly*, vol. 176, pp. 101–108, Jul. 1945. [Online]. Available: https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/

4. D. Engelbart and W. English, "A research center for augmenting human intellect," in *Proc. Fall Joint Comput. Conf.*, San Francisco, CA, USA: AFIPS, 1968, pp. 395–410, doi: 10.1145/1476589.1476645. [Online]. Available: https://dougengelbart.org/content/view/140/

5. T. Nelson, "Complex information processing: A file structure for the complex, the changing and the indeterminate," in *Proc. ACM 20th Nat. Conf.*, Aug. 1965, pp. 84–100, doi: 10.1145/800197.806036.

6. T. H. Nelson, "The heart of connection: Hypermedia unified by transclusion," *Commun. ACM*, vol. 38, no. 8, pp. 31–33, Aug. 1995, doi: 10.1145/208344.208353.

7. T. Nelson, *Ted Nelson Demonstrates XanaduSpace*. (Sep. 6, 2008). YouTube. [Online Video]. Available: https://www.youtube.com/watch?v=En_2T7KH6RA&t=3s

8. D. Engelbart, "Toward augmenting the human intellect and boosting our collective IQ," *Commun. ACM*, vol. 38, no. 8, pp. 30–32, Aug. 1995, doi: 10.1145/208344.208352.

9. H. Berghel, "ChatGPT and AIChat epistemology," *Computer*, vol. 56, no. 5, pp. 130–137, May 2023, doi: 10.1109/MC.2023.3252379. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10109291

10. H. Berghel and D. Blank, "The world wide web," in *Proc. Adv. Comput.*, M. Zelkowitz, Ed., New York, NY, USA: Elesvier, 1999, vol. 48, pp. 179–218, doi: 10.1016/S0065-2458(08)60020-0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0065245808600200

11. D. Engelbart, San Francisco, CA, USA. *The Mother of All Demos. Demonstration Hosted by the ACM/IEEE-CS Fall Joint Computer Conference*. (Dec. 1968). [Online Video]. Available: https://www.youtube.com/watch?v=B6rKUf9DWRI

12. A. Fisher, "How Doug Engelbart pulled off the mother of all demos," *Wired*, Dec. 2018. [Online]. Available: https://www.wired.com/story/how-doug-engelbart-pulled-off-the-mother-of-all-demos/

13. H. Berghel, "Social media and the banality of (Online) crowds," *Computer*, vol. 55, no. 11, pp. 100–105, Nov. 2022, doi: 10.1109/MC.2022.3198128. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9928241

**HAL BERGHEL** is a professor of computer science at the University of Nevada, Las Vegas, Las Vegas, NV 89154 USA. Contact him at hlb@computer.org.