UC Santa Barbara

UC Santa Barbara Previously Published Works

Title

AI and the University as a Service

Permalink

https://escholarship.org/uc/item/763635n0

Journal

PMLA/Publications of the Modern Language Association of America, 139(3)

ISSN

0030-8129

Authors

Kirschenbaum, Matthew

Raley, Rita

Publication Date

2024-05-01

DOI

10.1632/s003081292400052x

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at https://creativecommons.org/licenses/by-nc-nd/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Theories and Methodologies

AI and the University as a Service

MATTHEW KIRSCHENBAUM AND RITA RALEY

Turns out metanarratives did not collapse—in form and function, they simply moved to the drop-down menu. As with document templates in *Word* or *PowerPoint*, metanarratives of the present moment are modular, preprogrammed by software. But they are also propagated through public discourse by market-researched content, promoted through search rankings by optimization strategies, and promulgated through the social media feeds blasted to our devices by algorithmic amplification. For a timely example, one need look no further than the discourse around generative artificial intelligence. It is by this point commonplace to observe that something called "AI" has infiltrated, to some degree, nearly all domains of cultural production, scientific research, and economic activity. So rapid and so thorough has been the implementation that "AI" has itself become commonplace, and particularly so in the global context of telecommunications. We may accept or ignore the operations of such tools as spam filters, autocomplete, and recommendation algorithms, and we may remark on them with irritation and bemusement. But the tools themselves are now so embedded in everyday life, literally at hand, that they go without saying—and often without naming. They are regarded as customizable feature sets rather than the machine learning applications that they are. What better marker of habituation than the dissolution of the charged synecdoche "AI" into an array of functions, processes, instruments, and brand names?[1]

Nonetheless, "AI" retains an affective potency that can be mobilized on demand, whether for risk assessments or investment opportunities. This dynamic too has become commonplace. The sensational release of a new model activates both amazement and anxiety as we collectively register the shock of what has become possible and what is likely to come.[2] So too each high-profile error and injurious use (a chat application encourages monarchic

MATTHEW KIRSCHENBAUM is distinguished university professor in the Department of English at the University of Maryland, College Park. He serves on the MLA-CCCC Joint Task Force on Writing and AI and is completing a short book entitled *Textpocalypse*.

RITA RALEY is professor of English at the University of California, Santa Barbara. She is writing about cross-lingual language models and, with Minh Hua, completing her third piece on *GPT-2*.

© 2024 The Author(s). Published by Cambridge University Press on behalf of Modern Language Association of America
PMLA 139.3 (2024), doi:10.1632/S003081292400052X

assassination; a new image generation model goes predictably awry) animate dark fears of dystopian futures, as well as anger with corporations and institutions for having failed, again, to secure the public interest, to say nothing of basic humanity, which is perceived to be under threat from experimental nonhuman agents with evolutionary capabilities and the capacity for self-redefinition. Throughout, the AI model demos—as design fictions in every sense—do the work of conjuring up fantastic visions of profit and limitless growth.

Meanwhile, educational institutions respond by equal measures of pattern language, convening summits, commissioning white papers, charging ad hoc committees and task forces, and retaining executive consultants—themselves leading instigators of the very media cycles to which the institutions strain to respond.[3] That the output and outcomes of all this activity take the form of reports, visualizations, and slide decks—outputs that now can easily be machine-generated and by this point likely often are—is not so much symptom of a root cause but proof of concept of a continuous organizational strategy simultaneously anathema to the work of the university while scarily sanguine about the prospects for bringing the medieval institution into line with the data-drawn vectors of twenty-first-century capital.

We acknowledge, of course, that not all institutional responses are necessarily cynical or futile and that the possibilities for meaningful, affirmative, and even redemptive work have not entirely been foreclosed; if we thought they were, we would be writing a different kind of essay. Notwithstanding, what we have to offer is not a narrative of consolation, restoration, or repair. This is not another rallying call for the embrace of critique or resistance or retreat. Nor do we here enthusiastically advocate for the virtues of "leaning in"; the cultivation of new literacies; the opportunities afforded by, well, opportunity; or the intellectual rewards of interdisciplinary collaboration with data and computer scientist colleagues. Certainly, all of these are possible arguments, and all of them are necessary—indeed, we have each made them ourselves, repeatedly, and we will again in other contexts. But these arguments, themselves templated and, we suspect, less read than circulated, are not in our view adequate to the present moment for the profession's journal of record.

At issue is not only or perhaps even primarily "AI" as such. It is rather the all-too-common incapacity of educational institutions, of all types and at all levels, as well as their attendant professions, to be meaningfully responsive to the sociotechnical situation—that is to say, responsive in a manner other than denial, prohibition, capitalization, and capitulation. The rejoinder here might be that the deluge of summits and white papers suggests that perhaps from within the larger complex of partisan oversight, fiscal austerity, and brand maintenance an inspiring vision might emerge, however belatedly. One can always hope, of course, and even try to contribute to the effort. But there is a probability calculation to be made: Will the project devolve into an exercise in rationalization and containment, destined to be out-of-date before it even commences? Indeed, if the purpose is strictly managing the future, unsurprisingly, there are models for that. One can make allowances for, and even regard with some appreciation, the flurry of bans and blue books in the wake of the release of *ChatGPT* in late November 2022; educators are after all in the business of accreditation and were told, with some insistence, that the student essay was dead and the only way to preserve the academic mission was to cut the Wi-Fi. One also has to appreciate all the sacrificial labor that went into the educational town halls, tutorials, and guidance documents; for campuses (and bodies) still weathering the shocks and aftershocks of the pandemic, when no one has spare bandwidth for much else, a litany of well-intentioned "best practices" has been the default formula.

But shouldn't we—and here our enunciative "we" expands to the whole of the MLA, past and present—be able to do better? No more than a cursory scan of the contents of this journal since its inception in 1884 is needed to substantiate the basic point: our profession has all along been preparing its members to think and work through structural transformations in language practices.

For all the disparate methods and materials, for all the agonistic disputes, there has been a coherent project on the back end: we're thrown into this game called *language*, so let's figure out where it comes from, what it can do, what can be done with it, and how it might be changing. Mentions of actual technologies (typewriter, teletype, film, word processor) are sporadic up to a historical tipping point, and often simply with respect to office communications, but it is nonetheless also clear that, over the centuries, *language* and *literature* both have been understood to be subject to external forces that are, if not precisely technological, then at least material. Just as media environments have evolved, so too have scholarly and pedagogic practices, and we are now, as we have continuously been, exhorted to keep up, or to hold the line, to stabilize in relation to inherent instability.

Even so, for all the implicit and explicit worries over time that the profession and its objects are perhaps a little too susceptible to change—that the academy has lost its historical moorings or, now, that market share, mind share, and cultural influence are eroding—the prima facie assumption has been that the thing itself, *language*, was secure. Fierce contestations of practical uses and metaphysical questions might have suggested different rules for the game being played, but the symbolic ground seemed to remain constant, and intact. That symbolic ground, made up of human-constructed sign systems, can no longer be presumed a priori, and empirical study would confirm the argument. Such has always been the general condition of language in computational environments, of course: we "don't know what our writing does," because when we type, the flickering signifiers that appear on the screen mask a cascading series of translations from character encoding down through levels of programming languages and assembly code, at the root of which is machine code's manipulation of electrical voltages. Input and output signals, in other words, have been undergirding language as such for many decades of this journal's publication.

However opaque, and however inaccessible to human perception, these processes of translation, numerical representation, compilation, and execution seem now, in the wake of almost-dizzying advances in the domain of natural language processing (NLP), to be quite simple, and quaintly so. We may not have known *what* our writing did, either technically or theoretically, but in actual practice the scene of writing was unambiguous: press the keys, and the words you typed were (generally) the words you got. The advent of large language models (LLMs) has radically transformed this technolinguistic situation, full stop.[4] More precisely, the implications of the already-widespread and now-accelerating implementation of LLMs are both epistemic and epistemological. And, more concretely, the break has been initiated by the entire end-to-end apparatus of machine learning for language processing—from the initial formation of massive unlabeled training datasets and subsequent tokenization and word embeddings, through complex model architectures (GPT, BERT) and model training, up to an array of sampling techniques (top-k, top-p) and postprocessing steps (filtering, personalization). Thus the scene of writing in this historical moment: prompt a model, and the words you see have been probabilistically generated based on the extrapolation of linguistic patterns from training data composed of sequences of tokens that have been converted into numerical representations.

"Prompting a model" is, appropriately, both a specific activity and an abstraction, just as a "model" is itself both an actual instrument and an epistemic thing. Prompting—calling a model to respond, "chameleon-like," to user input, as OpenAI explained when it partially released *GPT-2* in February 2019 ("Better Language Models")—is one of a series of cultural techniques for conditioning language model behavior, in addition to fine-tuning, metrics, and benchmarks. But the phrase "prompt a model" also functions as shorthand or a simplified representation or perhaps again a synecdoche for the practice of engaging an increasingly extensive ensemble of applications and interfaces built on top of, or otherwise integrated with, an equally extensive ensemble of language models, large and small, proprietary and open. Put more plainly, prompting a model already means a good deal more than typing

words into a text-entry window. So-called prompt engineering, an inexact science to be sure, takes as its core mission the design and development of methods by which to instruct, cue, or otherwise initiate the process of language generation so as to result in "good" output. What constitutes the good, and how it is identified and evaluated, are matters to which we return below (and here we revert to the authorial "we").

It is difficult to avoid further mention of *ChatGPT*, although we would prefer it if we could because the term, perhaps as a consequence of insistent repetition, has come to mean nothing, something, and everything, all at once. For all its semantic ambiguity and imprecision, the term *AI* does at least summon a speculative, ordering imagination that sweeps the ethicopolitical into the domain of science and informs discussions of bias, fairness, mental health, trust, safety, open systems, a data commons, and public resources. In contrast, the term *ChatGPT*, precisely because of its common use as shorthand for the broader sociotechnical condition we have been sketching, has the effect of short-circuiting thought itself—it is referential and self-referential, all at once. Practically, *ChatGPT* refers to a chat application built on top of an LLM (the Generative Pretrained Transformer), the technocultural authority of which derives in part from its accessibility and ease of use. A chatbot's end-user interface is indeed more accommodating than OpenAI's "playground," which allows users to control hyperparameters such as temperature, frequency penalty, and max tokens (even using language such as this risks losing a mass audience; hence the appeal of an interface with only a basic messaging window and familiar menu options). But there already have been, and will be, many more chat applications and application programming interfaces integrated with, or otherwise supported by, many more LLMs, all of which will close off access in the name of granting access. Indeed, *ChatGPT's* position at the top of the proverbial leaderboard has already started to give way to other models and applications, and there are a number of scenarios in which it simply disappears.[5] Nonetheless, the term *ChatGPT* continues to index a host of objects, agents, activities, behaviors, and attitudes that extend well beyond the strictly technical domain.

We return then to the question of how the profession can best position itself within this milieu. On the surface, the new scene of writing—that is, language processing—may seem to be composed simply of permutations of familiar activities, processes, and concepts: symbolic representation, semantic relationships, corpus construction, translation, indexicality, the evaluation of output, authorship, and generic forms. In this version of the class, there is at least still a text. So too with word embedding, the computational representation of words with numbers—more precisely, the representation of words in a continuous vector space by assigning them numerical values, or coordinates—which is fundamental to computational text analysis as it has been historically leveraged within the digital humanities.

What is less familiar about the ecosystem of LLMs, and not often remarked in a context such as this, is subword tokenization, one of the crucial NLP techniques. Subwords are as they sound: not words but rather parts of words that bear no necessary relation to the linguistic units identified and categorized by human sciences. They can be morphemes, but not necessarily or exclusively. A good example is *th*: statistical analysis of any English-language corpus would probably capture the frequency of this pairing and identify it as a subword token, by means of Byte Pair Encoding. Digraphs are not of course unrelated to writing as it has historically been understood—after all, morphological rather than linguistic principles can inform the typographic element of the ligature. But the ligature is only about appearance, and it presents no challenge to the operative principles of language. Subword tokens, in contrast, result from a process of segmenting words within a corpus according to statistical or rules-based criteria. These segmented tokens are then used to construct what are called vocabularies (that is, collections of parts of words) that help a model learn the statistical representation of text data. And here the names of tokenization algorithms and libraries such as

WordPiece and SentencePiece are particularly suggestive. Perhaps linguistic protocols have been eclipsed by arithmetic means. Perhaps the leading sciences and technologies are no longer operated and mediated by language *as such*. And perhaps we truly are at the beginning of something new.

At the 2024 MLA convention, "Reading Generative AI: Theory, Data, Critique" was one of three pilot sessions convened to test a new format: professional development seminars "meant to catalyze scholarship on research topics that cross areas of study, disciplines, methods, and perspectives" ("MLA Convention Seminars"). After submitting abstracts the preceding spring, a dozen participants were selected. They had proposed work on a wide variety of questions and problems related to the topic at hand and represented a range of institutions, approaches, and career stages. Attention then turned toward individual and collective preparation for the seminar, which included many months of drafting and peer review. Once in a room together in Philadelphia, we engaged in a spirited three-hour discussion, one whose breadth, depth, and vibrancy provided the impetus for the cluster of essays that appear here.

Of course this was not the only conversation about AI at the convention: indeed, there were enough dedicated panels and individual papers that it would not have been possible to attend them all. But we recap these details about the professional development seminar because neither its mode nor its deployment at the convention is incidental to the broad sketch we offer in the preceding paragraphs. The three initial topics—generative AI, prison literature, and the medical humanities—all suggest new directions for research, teaching, and even activism. The seminar format itself is an experiment occasioned in part by the new realities of the MLA convention: with the displacement of scant hiring activity to remote platforms, as well as the staggered recovery from the pandemic and the reduction in travel support for nearly all its constituents, the convention seems clearly to be seeking new ways to position itself as the chassis for the profession(s).

This journal is another such chassis. In the context of a previous Theories and Methodologies section on distant reading and computational approaches to literary studies (May 2017), a language model would have been understood as a statistical profile of a text or corpus of texts. This description holds true for LLMs as well, but the scales are vastly different; *GPT-4*, for example, reportedly consists of 1.76 trillion parameters (that is, variables that determine the model's behavior and capabilities). What this indicates is not yet another novel computational application or approach but rather a general condition of language and life (or, again, an episteme). If some of the public enthusiasm around digital humanities has cooled in recent years, it is perhaps because it has been a victim of its own success, at least insofar as the kinds of "theories and methodologies" previously singled out for special attention are increasingly regarded as altogether ordinary, if not normative, within the disciplines and subdisciplines they serve. By contrast, whatever uses LLMs may prove to have within domain-specific literary and historical research will be well downstream of the far more basic challenges posed by the technology and its market uptake, yielding institutional configurations that may hold out little promise of exceptional status for even the most digitally forward humanists.

Some sense of urgency about the emerging situation is evident in the fact that just two months after the public release of *ChatGPT*, the MLA moved to partner with a peer organization to convene the MLA-CCCC Joint Task Force on Writing and AI. The charge included "[t]aking stock of the current state of the issue and identifying implications for teachers, students, organizations, and scholars."[6] Since January 2023, the task force has released two (soon to be three) public working papers, offered multiple webinars, and curated a selection of instructional materials. All this activity is attuned to the exigencies of the moment, and yet it also—and precisely *not* paradoxically—conforms deeply to scripted organizational norms. At the same time, the task force is not a committee or standing entity of either

organization, which means that its working papers (note the provisional and amorphous form), while sponsored and reviewed by the two executive boards, do not carry the weight of official policy. Thus, the agility demanded by the moment also begets a deferral of agency and responsibility.

The task force itself, we wish to be clear, has done vital and necessary work (one of us, Matthew Kirschenbaum, is a member); but like a seminar or a journal feature, the mode is symptomatic. However critical or otherwise aligned with the profession's values and expressed commitments, the responses to AI largely subsist within a bounded and belated repertoire, a circumstance from which, we hope we have made clear, we do not seek to exclude ourselves. There has been an outpouring of enterprising activity across the institution of the university, all of it (one imagines) capturable by the metrics of performance review. That there is actual and symbolic profit to be made (in whatever small measures of reputational currency, honoraria, and grant monies) is not incidental or too unbecoming to acknowledge. But transactions on the level of the individual are magnified in larger disciplinary economies as we all conceive new curricula and assignments, new certificates and degree programs, all straining to meet the moment by demonstrating "leadership" once again.

With that as a substrate, we now relate, in highly stylized terms, the principal disciplinary operations functioning in correlation with LLMs and generative AI. We adopt as our own rhetorical device the model, figuratively speaking—in fact a suite of figurative models fine-tuned (as it were) from our disciplinary foundation models. The foundation model itself, we should say, is the true heir to the metanarrative, being the de facto prescriptive, closely guarded, and contested core from which commercial downstream models are spawned for individuated tasks and applications. Here then we put forward some imperfectly rendered, inevitably nondiscrete, and appropriately idiosyncratic models of models, presenting in each instance first the foundation model and then, after the arrow, its fine-tuned elements, followed by a short gloss. Whether, how, and to what extent they may prove sustainable through the institutional transformations we outline in the conclusion of this essay must remain an open question.

Critique → Historicism, Political Economy, and Materialism

The deep and abiding commitment, even faith, in the potential of historicization and attention to conditions of production. The pursuit of grit, of ghosts in the machine, and the insistence of and on the real, the lived, the situated, and the embodied, whether wageworkers cleaning data for pennies on the hour or the environmental harms of resource extraction. The enviable certainty that this kind of documentary, even forensic exposure will effect meaningful structural transformation, or if not that, then "better" models of consumption or principled refusal. Surely materiality must (still) matter! But also the inevitably overdetermined lines of inquiry that begin and end with the pronouncement that these tools and technologies are overblown, unnecessary, grossly commercial, in fact unprofitable and thus unsustainable. They may be all that, but the stronger move for humanities critique now, given the already deep-rooted monopolistic hold on both the research and the resources, is to enact what theoretical scaffolds we can on top of new technical infrastructures and think from there.

Engagement → Pedagogy and Practice

The work, no less essential than that of critique, of figuring out what to actually do, and first in our most publicly relatable habitus, the classroom. Unsurprisingly, the administrative and pedagogical burdens of both reinvention and circumnavigation fall disproportionately on those who may not have the luxury of engaging otherwise: not only those with obscene course loads but also instructors who do not set their own syllabi, writing center tutors, and the staffers in teaching resource centers. But even if granted all the time and resources to imagine different possibilities, with a full suite of

workshops and guidebooks and teach-ins, it is not going to be feasible to devise a strategy or method that can be ported to all institutional situations, or even to all assignments. For an inherently fluid technological situation, with new applications and implementations seemingly by the day, and the models themselves continuously evolving, there can be no stable and singular path forward. Of course the motivational clichés, symbolic rewards, and so-called incentive structures that often seem to accompany such opportunities to innovate are cynical and nakedly transparent; but the posture of engagement is not to be dismissed, even if we had the luxury or the choice. The work is too real, if not strictly of our own choosing.

Data Work → Origins and Archives, the New Data Philology

In September 2023 *The Atlantic* published a database indexing nearly 200,000 books in the so-called Books3 corpus (a massive garbage patch of raw text data widely believed to have been used to help train a number of foundation models), thereby provoking a firestorm of outrage as one by one authors queried and discovered their own published words scooped and scraped into its maw ("These 183,000 Books"). However justifiable the anger, can we not remark too on the reinvestment in origins, the earnest searching for sources that attends the new copyright fundamentalism? For surely, deep down in the muck at the very bottom of the model, there must still be a discrete text to be identified, much less interpreted—mustn't there? But framing the privatization of creative work in traditional, dare we say conservative, terms of authorship and intellectual property then necessitates a concession to a market economy for culture. Protracted, expensive, and often unpredictable litigation may be unavoidable, although the matter at hand is not simply property theft: it is rather an enclosure of the language commons. Let us then instead work to cultivate and support alternative data resources. Let there be legibility all the way down. Open and transparent training data, as well as models, means accountability. The widely used community platform *Hugging Face* is perhaps the marquee effort here, given its extensive repository of open-source models and datasets, but it is still a privately held entity. To underscore the broader point: there is still an acute need for public data hubs, to which the humanities can bring to bear its rich traditions of curation and stewardship. A new *HathiTrust*, but explicitly for the language model ecosystem.

Experimentation → New Methods, Embodied Forms of Knowledge

The spirit of critical making, DIY culture, hacker manifestos, the creative avant-garde, and open-source activism lives on in the motivated and reflexive use of new technical methods. As with most technological systems, computing not least, there are all manner of research questions, theoretical insights, and aesthetic categories to be explored through open-ended, practice-based research. How might machine learning systems be made more legible, fair, or even exploited? Myriad artistic and literary explorations fall within the purview of experimentation: we may not be able to pretrain large models, but we can at least content ourselves with fine-tuning, postprocessing, and inference. Alternatively, we can build small, bespoke datasets. Or we can go the other route and adopt adversarial techniques such as cloaking, perturbation, poisoning, and more general practices of data manipulation. We are experimenting rather than being experimented on! The API-ification of AI—in fact the wholesale transformation of AI into a commodity through application programming interfaces that grant access to machine learning models through proprietary platforms—of course thwarts experimental practices, but, happily, someone always figures out how to navigate around the latest corporate model.

Evaluation → Aesthetics and Critical Judgment, Practical Criticism

"Bot or not?" It's an old question: resolved or abandoned by philosophy, now relegated to offices of academic integrity. "Let a thousand models bloom, and

a thousand detection systems bloom in turn." (We are twenty-seven percent certain the previous sentence was generated and ninety-nine percent certain this ship has sailed.) *Human* and *machine* were only ever prescriptive keywords, but maybe it's time to let go of the heuristic and accept that it's all entangled, all the way down. From there, all manner of possibilities and questions present themselves. What are our newer aesthetic categories? Are we sure we know what "good" output actually is? Can the basis for evaluation be concretely identified and scored, or is it rather a tacit hermeneutic contract (we know a good sentence when we see it)? The strong pivot toward LLM optimization, particularly as evinced by the launch of the "GPT Store," with its growing number of customized GPT models, would suggest a consensus view: the benchmarks now work well enough, so all that remains is application, which is to say monetization and personalization. This then is the time and space for intervention. Let us bring the entire tradition of critical judgment—as it has been honed, debated, and theorized over the centuries—to bear on the problem of how to qualitatively evaluate model output.

Professional-Managerial Work → New Templated Forms for Teaching and Research

The model we have already acknowledged by belaboring our own situation within a convention, an organization, and a journal, to say nothing of the departments and institutions that employ us. Today's cycles of just-in-time publications, working papers, and symposia will condition tomorrow's resource allocations and material infrastructures. As campus-level messaging suggests, there will be no escaping the latest realignment. Much of the actual work that follows—the implementation of everything from strategic visions to curricular transformation, workforce rationalization, and compliance regimens for updated security protocols—will be predetermined, even if some of it may prove salutary and intellectually rewarding. All of it, however, will be inseparable from the political (in the sense of political economy and partisan politics both), economic, and institutional circumstances of the university in ruins, about which we have more to say now, in our concluding paragraphs.

A video in our social media feeds. The setting is a crowded college classroom, the production values amateur. To a soundtrack of bouncy pop from the Romanian singer Inna, a young woman looking morose enters the frame at the front of the room and a captioned narrative begins: the woman, who is the professor, is "sad" because no one needs to take her classes anymore. In fact, no one needs to take anyone's classes. It's not the Internet—her students now don't even need *Google*. It's AI, or more specifically in this instance a language model being marketed as *Blackbox.AI*, which exhibits the ability to ingest a link to a *YouTube* clip and deliver a short prose summary. Students are learning not by watching videos—what a time sink!—but by reading brief narrative summations generated by the model, which does the tedious work of "watching" and synthesizing for them, as well as coding their assignments. The professor is a quick study and knows that the curtain on her role is coming down. All that remains to do is open her own *Blackbox.AI* account.

The video we are describing originates from an account created in April 2022.[7] It is a small-scale influencer, with only some six thousand followers, though what appears to be a linked or shadow account for distribution of the same content is an order of magnitude larger, at forty thousand followers. Both accounts (themselves likely automated to a large degree) host a rotating stable of videos that unfold according to the basic script we have just outlined. Worth underscoring is that both the institution of the university and the Internet of Google are presented as obsolete in equal measure. After all, typing a query into a search window, browsing through the results, and then assessing what you find there is *so* 2022. No more sifting and sorting is the promise, no more curating information, just autosummary on dialogic demand from a vast undifferentiated pool of content.

It is (of course) without a doubt snake oil or, as is said, "hype." It is also simply garden-variety NLP with a feature set based on video-to-text and some other multimodal capabilities. But what the

example of *Blackbox.AI's* promotional campaign demonstrates is the emergence of yet another kind of model or template, not a language model strictly speaking but a productivity schema laid across the full spectrum of the postindustrial knowledge economy. The clear, and templated, message is that the university no longer has any purpose at all: they came for our content first and shut the door behind them with our coding instruction. We thus arrive at last at the long-previewed moment when the map has gotten just about as big as the territories. Language—in the form of executable code, machine-readable data, and human-readable text subject to tokenization for word embeddings—is not merely product but also managerial apparatus (yes, *dispositif*) that produces humans as subjects of, and subjects to, the extractive economies that are the new scene of writing. Certainly the truth that some—many—geopolitically predisposed bodies will be consigned to the irreducibly real and physically devastating portions of the labor cycles that enable and sustain LLMs is not beside the point. But even those whose industries subsist within the domain of what was once aspirationally called "knowledge work" will be incorporated within those selfsame labor (and life) cycles, consolidated and controlled by the unregulated corporations whose own capital interest is the only and ultimate goal.

Meanwhile, glimmerings of language's new currency have revived optimism about a future in which these very corporations are eager to hire humanities students, a seemingly welcome respite from the relentless drumbeat of the death of our disciplines. The reality, however, is that it is precisely this tokenizing of language—both its subordination to technical processes and its symbolic devaluation—that promises to render universities vulnerable to the market logics by which the neoliberal institution has staked its primary claim to a sustainable, if not survivable, future. If Google coding certificates are beside the point, what chance have even the R1s? In this analysis, both *Blackbox.AI* and *ChatGPT* are epiphenomenal manifestations of an ongoing sequence of structural changes to both education and the workplace. Where, then, do we imagine the university in such a system? What outside can there be when the glue (or grease) of the machine is language itself, as much as can possibly be scraped for the foundation cores? (The culture industry—with its newly urgent questions about creativity—is our ground zero.) Again, we hasten to underscore that none of this is a function of claims about the technology or its capabilities *as such*, but rather about the way in which the technology is symptomatic of latent logics and transformations that have long been underway. Put plainly, if "AI" didn't exist it would be necessary to invent it.

Of course, the institution is also a market. And that may prove its salvation. Consider all the third-party software licensing for course management and evaluation, as well as the new entrepreneurial ventures that entice budget-strapped schools with the promise that they no longer need to worry about building and supporting a technological infrastructure. For these, a representative corporate pitch: because you lack both equipment and staff, we will remotely operate the cameras in your lecture hall so that you can meet the requirements of the Americans with Disabilities Act, and in exchange we are going to store those lectures and use them to train our in-house language models, access to which we will license to you in the next funding cycle, when you are ready to add our transcription services to your subscription. If this scenario seems far-fetched, recall the proposition of an "app" replacing language instructors at a flagship public university—a supposed efficiency measure floated by an executive administrator still extolled as a thought leader in American higher education ("WVU Plans").

In this, our own sector of the economy, we can now begin to apprehend the true and actual foundation model subsuming the disciplinary ones we were at such pains to enumerate above: the university not in ruins but *as a service*. The idea of the University as a Service (UaaS) extends the model of Software as a Service (SaaS) to education: physical institutions (for now) provide the lecturers, content, and degrees; in turn, the technological infrastructure, instructional delivery, and support

services are all outsourced to third-party vendors and digital platforms.[8] Licensing and subscription agreements favor short-term budget planning; so too do they enable an administrative vision of universities as customizable, scalable, cost effective, and available on demand. And thus the decision-making is too often exclusively in the hands of CIOs, IT staff members, and instructional development, with academic affairs relegated to the position of managing the implementation of commercial ed-tech applications that promise continuous pedagogic improvement, which is now to be accelerated by new AI features, all of them generating revenue through the generation of data.

Those of us in what positions of authority remain—holding on to faculty governance offices, seated on the relevant committees, presumably able to transmit messages to the right sets of ears—may try to forestall these developments. Every outside vendor contract is still a negotiation, literally and otherwise. But it is unlikely to be that simple. The firewalls are coming down, and the doors of the institution are open, not only to lifelong learners and citizen scientists or even just the start-up hucksters, but to the flash mobs mobilized in the culture war that is also a cold civil war. If it is true that the future belongs to crowds, it seems clear by now it will be less a valorous multitude than the malign cults of personality that accrete around billionaires, politicians, billionaire-politicians, celebrities, and other influencers, all targeted and manipulated by the outrage merchants who command the largest and most lethal followings on social media. And so the new Trojan horse: to the extent LLMs continue to be trained on the torrents of platforms like *4Chan*, *Reddit*, and *X*, the positions and agendas expressed therein will be imported directly into future foundation models. Those foundation models will then be iterated, localized, and branded—complete with mascot imagery—before being sold off to individual institutions as boutique products vertically integrated with campus services, from marketing and communications to health and public safety.[9] What this looks like for the several large publics currently piloting or plotting exactly such services may initially appear relatively benign; we wonder what it will look like at the growing list of places already overtly subject to ideological capture, the systems and campuses that find themselves deep in the red, whether politically, financially, or, as is often the case, both. Mission alignment as fine-tuning as language-learning-*thinking* optimization.[10]

We do not pretend to have the answers to the very real questions that have begun to infiltrate these final paragraphs. We have tried to articulate the gravity of what we believe to be at stake and we invite readers of this journal to think with us, as we have invited the participants in our seminar to do for the essays that follow. What to do with the remains of the day? What comes after even the ruins have been repossessed and enclosed? These too are real questions. Some will no doubt wish to declare crisis bankruptcy. We do not begrudge them—crisis is its own industry, and we are not crisis managers. But this much we will say: AI is not the prompt, it is the punctuation.

Notes

All datasets are necessarily incomplete, but the following named entities have informed the text we have generated here (readers may recognize allusions and common multiword expressions): Louise Amoore, Claudia Aradau, Hannes Bajohr, Jean Baudrillard, Emily Bender, Walter Benjamin, David Berry, Lillian-Yvonne Bertram, Mercedes Bunz, Joy Buolamwini, Rüdiger Campe, John Cayley, Michel de Certeau, Wendy Hui Kyong Chun, Sarah Ciston, Kate Crawford, Nan Da, Gilles Deleuze, Ranjodh Singh Dhaliwal, Wai Chee Dimock, Stephanie Dinkins, M. Beatrice Fazi, Mark Fisher, Michel Foucault, Seb Franklin, Alex Galloway, Timnit Gebru, William Gibson, Lisa Gitelman, David Golumbia, Félix Guattari, John Guillory, Orit Halpern, Michael Hardt, Adam Harvey, N. Katherine Hayles, Leah Henrickson, Minh Hua, Patrick Jagoda, Vladan Joler, Brian Justie, Frederic Kaplan, Christopher Kelty, Friedrich Kittler, Kari Kraus, Jonathan Lethem, Alan Liu, Jean-François Lyotard, Adrian Mackenzie, Angelina McMillan Major, Albert Meroño-Peñuela, Colin Milburn, Philip Mirowski, Margaret Mitchell, Antonio Negri, Sianne Ngai, Fabian Offert, Trevor Paglen, Luciana Parisi, Allison Parrish, Everest Pipkin, Bill Readings, Jennifer Rhee, Anna Ridler, Jonathan Roberge, Russell Samolsky, Bernhard Siegert, Hito Steyerl, Lucy Suchman, Eugene Thacker, uncertain commons, Ted Underwood, McKenzie Wark, Leif Weatherby, Raymond Williams, and Shoshana Zuboff.

1. In keeping with the arguments we are making about the impossibility of externalization from the new systems of linguistic totalization, we have enlisted certain keywords from the fields of artificial intelligence and natural language processing—including *model*, *foundation model*, *prompt*, *fine-tuning*, and *token*—to serve as both technical terms of art and conceptual scaffolding. For the benefit of readers for whom this terminology may be unfamiliar we offer the following brief explanations. *Natural language processing* names a decades-old applied research area at the intersection of linguistics and computer science dedicated to using computers to analyze as input and produce as output human speech or text. Both noun and verb, *model* is a particularly important word for us: we use it in its instrumental sense to refer to the statistical profiles of data corpora that form the basis of machine learning applications, but also in the long-standing and generic sense of a representational abstraction of a system, a process, or an entity. It is in this latter capacity that the second section of our essay enumerates some knowingly compressed and inevitably situated "models" of current academic and institutional practices around AI. A *foundation model* is a large, pretrained model used as the basis for developing specialized, "downstream" models tailored for domain-specific tasks, a process known as *fine-tuning*. As their allusive character invites, we also use both *foundation model* and *fine-tuning* in relation to the aforementioned "models" of practice and discourse. A *prompt* is an incitement to action or speech, but it is also the currently accepted term for initiating an interaction with a machine learning model. Finally, as we explain in more detail in the essay, a *token* is both an individual unit of text (a word, subword, or character) that has been extracted from a larger corpus and a unit of value or exchange. While we trust that context and good judgment will serve to disambiguate our use of each of these terms in their actual versus more figurative or allusive dimensions, their overlap and slippage is salutary as a demonstration of the ever-increasing imbrication of language, statistical mathematics, and economics that is the epistemic frame of this essay.

2. The February 2024 release of OpenAI's text-to-video model Sora conformed to this template, as did the May 2024 release of *GPT-4o*, mere days before we finished the copyedits on this essay.

3. A suggestive parallax view of the university's response to contemporary developments in AI can be obtained through a historical comparison with the transformative developments in biomedical research in the late twentieth and early twenty-first centuries: attendant upon advances in stem cell research and genomics was an explosion of interventions (biomarkers, accounting systems, trials, ethical protocols, IP regimes), all managed by a class of new professionals (evaluators, principle investigators, ethicists, advocates). Whether the process of instrumentalization, leveraging, and capitalization now underway is a fundamental paradigm shift or merely accelerative is a question still to be resolved. So too is the question of whether the university will have forgotten how to assert ownership of its own research.

4. With this claim we must necessarily acknowledge the centrality of the English language for NLP. Why and how this came to be might be intuited, but the future is less certain. Further development of LLMs for Chinese, Korean, Spanish, French, German, Arabic, Russian, and Japanese will perhaps mitigate this skewed representation; more important will be the development of multilingual corpora and cross-lingual models, the standardization of character encoding, and support for so-called low-resource languages.

5. For example, as of our writing in early March 2024, *Opus* (the largest version of Anthropic's model, *Claude* 3) is outperforming *GPT-4* across a range of NLP benchmarks.

6. See "MLA-CCCC" for the working papers and other materials produced by the task force.

7. See Promotional video for *Blackbox.AI*.

8. Foremost among these are *Google Workspace*, one of the first major enterprise services, not incidentally tested at a city college; the course management system *Canvas*; and, of course, *Zoom*. Not only are these services all integrated with one another and with myriad others (among them course evaluation software and transcription tools) but they all now, as expected, integrate AI technologies so that the UaaS might circumvent upstart platforms such as *Blackbox.AI*.

9. The University of Michigan announced a suite of UM branded custom services based on GPT and other models in August 2023; in January 2024 Arizona State University announced a partnership with OpenAI, including the use of personalized applications for writing instruction. Since then the gates have opened: *ZotGPT* at the University of California, Irvine, *TritonGPT* at UC San Diego, and doubtless many more after the rollout of OpenAI's "responsible" design for universities: *ChatGPT Edu*. And so the future arrives, unevenly distributed as always.

10. This then is the place for a necessary date stamp inserted at time of copyedit. Having completed this essay in early March 2024, before the advent of the widespread, tumultuous, and still-unfolding campus protests, we are all too aware that now, some months later, we risk instant obsolescence (or worse, glaring naivete) because whatever harms will follow from "AI" might seem to have been eclipsed by the events of the spring. But these phenomena are not distinct; rather, they are coterminous and cognate, as (in far too many cases) templated administrative responses demonstrate. Just as virtual infrastructure is outsourced to vendors and contractors, so too is policing and physical security, as underscored by the sudden appearance on quads and greens of entities like Apex and CSC to provide just-in-time defense operations, complete with body armor, drag-and-drop bollards, and tear gas. Above we claim that AI is "symptomatic" of "latent logics and transformations that have long been underway" and that if "AI didn't exist it would be necessary to invent it," and here we can extend the argument that this is all of a piece with the same complex of thinking by pointing to the UaaS scripts that were field-tested during COVID-19 and re-executed for the encampments. The "service" the university provides in this time of crisis is not just that of its data (though much data are generated through added securitization) but also that of a field of deployment, or *platform* in the sense that the physical site of the campus is arrogated as a nexus for the capitalized ideological vectors that ramify throughout algorithms and "services" of all kinds, of which

the service named higher education is now merely one homologous epiphenomenon.

Works Cited

"Better Language Models and Their Implications." *OpenAI*, 14 Feb. 2019, openai.com/index/better-language-models/.

"MLA-CCCC Joint Task Force on Writing and AI." *Humanities Commons*, 2024, aiandwriting.hcommons.org/.

"MLA Convention Seminars." *Modern Language Association*, 2024, www.mla.org/Events/MLA-Convention-Seminars.

Promotional video for *Blackbox.AI. Facebook Reels*, uploaded by The IG Goat, www.facebook.com/reel/378597761437258. Accessed 3 Feb. 2024.

"These 183,000 Books Are Fueling the Biggest Fight in Publishing and Tech." *The Atlantic*, 25 Sept. 2023, www.theatlantic.com/technology/archive/2023/09/books3-database-generative-ai-training-copyright-infringement/675363/.

"WVU Plans to Completely Cut World Languages Department." *Language Magazine*, 15 Aug. 2023, www.languagemagazine.com/2023/08/15/wvu-announces-plan-to-dissolve-entire-department-of-world-languages-literatures-and-linguistics/.