Report of the Yale Task Force on Artificial Intelligence

Reflections and Recommendations

JUNE 2024

Table of Contents

Yale Task Force on Artificial Intelligence Members 1

Report of the Yale Task Force on Artificial Intelligence 2

- Introduction 2
- Themes and Observations from University AI Panels 3
- Yale's Unique Strengths 5
- Recommendations for Investments and Coordinated Action 5
- Conclusion 7
- Acknowledgments 8
- Appendices 8
  - *Appendix 1: AI Summary—Education* 10
  - *Appendix 2: AI Summary—Yale School of Architecture* 16
  - *Appendix 3: AI Summary—Yale School of Art* 20
  - *Appendix 4: AI Summary—Yale Divinity School* 23
  - *Appendix 5: AI Summary—David Geffen School of Drama at Yale/Yale Repertory Theatre* 26
  - *Appendix 6: AI Summary—Yale School of Engineering & Applied Science* 28
  - *Appendix 7: AI Summary—Yale School of the Environment* 33
  - *Appendix 8: AI Summary—Faculty of Arts and Sciences: Humanities* 36
  - *Appendix 9: AI Summary—Faculty of Arts and Sciences: Basic Sciences* 38
  - *Appendix 10: AI Summary—Faculty of Arts and Sciences: Social Science* 41
  - *Appendix 11: AI Summary—Yale Jackson School of Global Affairs* 46
  - *Appendix 12: AI Summary—Yale Law School* 58
  - *Appendix 13: AI Summary—Yale School of Management* 63
  - *Appendix 14: AI Summary—Yale School of Medicine* 71
  - *Appendix 15: AI Summary—Yale School of Music* 79
  - *Appendix 16: AI Summary—Yale School of Nursing* 82
  - *Appendix 17: AI Summary—Yale School of Public Health* 90
  - *Appendix 18: AI Summary—Yale College* 96
  - *Appendix 19: AI Summary—Collections and Scholarly Communication* 105
  - *Appendix 20: AI Summary—Clinical Practice* 109
  - *Appendix 21: AI Summary—Operations* 115

Yale Task Force on AI Members

Jack Balkin, Knight Professor of Law and Professor in the Institution for Social and Policy Studies

John Barden, Vice President for Technology & Campus Services

Phillip Bernstein, Professor Adjunct, Yale School of Architecture

Michael Crair, Vice Provost for Research, William Ziegler III Professor of Neuroscience, and Professor of Ophthalmology and Visual Science

Julie Dorsey, Frederick W. Beinecke Professor of Computer Science

Jennifer Frederick, Associate Provost for Academic Initiatives

Marla Geha, Professor of Astronomy and of Physics

Susan Gibbons, Vice Provost for Collections & Scholarly Communication

Murat Günel, Sterling Professor of Neurosurgery and Professor of Genetics and of Neuroscience

Trace Kershaw, Susan Dwight Bliss Professor of Public Health (Social and Behavioral Sciences)

Chad Losee, Head of Strategy, Office of the Provost

Lucila Ohno-Machado, Waldemar von Zedtwitz Professor of Medicine and Biomedical Informatics and Data Science

Nilakshi Parndigamage, Vice Provost for Strategic Initiatives

L.A. Paul, Millstone Family Professor of Philosophy and Professor of Cognitive Science

Phillip Atiba Solomon, Carl I. Hovland Professor of African American Studies and Professor of Psychology

Scott Strobel (chair), Provost and Henry Ford II Professor of Molecular Biophysics and Biochemistry

John Tsang, Professor of Immunobiology and Biomedical Engineering

Nisheeth Vishnoi, A. Bartlett Giamatti Professor of Computer Science

Introduction

Background

Artificial intelligence (AI) has existed for decades, though its recent transformation into a readily available tool has raised significant questions about its role in society. Today, individuals from a wide range of industries, professions, and backgrounds are exploring what it means to work, discover, think, solve, produce, and create in an age where rapid technological shifts change the nature of our engagement with these very human-centric endeavors.

As AI continues to develop and proliferate, the university—an institution rooted in the production and dissemination of knowledge—is uniquely positioned to lead technological advancements as well as conversations about their human implications. Yale University's mission to "improv[e] the world today and for future generations through outstanding research and scholarship, education, preservation, and practice" serves as inspiration for it to develop AI and formulate scientific, technical, ethical, legal, and social perspectives that will shape how individuals consider, harness, and engage with this technology. Combining the university's capacity for cutting-edge research and innovation with humanistic and artistic inquiry, Yale is exceptionally well equipped to simultaneously advance foundational AI and its applications and shape global considerations for its ethical and responsible use. Additionally, as one of the world's leading educational institutions, Yale is particularly qualified to train global leaders who will use AI in ways not yet imagined and establish frameworks for how society engages with this technology in the future.

In fall 2023, in recognition of the opportunities described above, Provost Scott Strobel asked Yale's deans to explore AI within their communities. In particular, he asked that deans and their faculty members consider and define an approach for addressing the opportunities, promises, and perils of AI for their respective fields. In response to this request, each dean convened a panel, or in some cases panels, of faculty who engage with AI. These panels, composed of researchers and scholars from a range of disciplines, backgrounds, and approaches, provided perspectives on emerging AI advances and critical areas of inquiry, with a focus on AI as a tool for academic research, a technology to be innovated upon and improved, and an ethical and societal phenomenon to understand and shape. In addition to these dean-led efforts, university leaders in education, collections, clinical practice, and operations convened their own panels of Yale experts to reflect on the implications of AI in these areas.

As a result of convening, the schools and units produced written summaries of their efforts and goals related to AI. These summaries (attached herewith as appendices), provide direction and strategy for each school and unit, recognizing that the expertise these communities will contribute to AI development, policy, and practice will differ. These school- and unit-level plans are critical given that AI will impact different fields in fundamentally different ways.

The Role of the Yale Task Force on AI

In January 2024, in parallel with the activities of dean- and unit-led panels, Provost Strobel assembled our Task Force — a wide-ranging group of faculty and senior leaders from across the university — to review AI activity through a university-wide lens. This effort, supplemental to school- and unit-specific work, recognized that for Yale to be at the forefront of inventing, informing, and developing AI — in individual disciplines and across domains — the university must draw on expertise from the full breadth of its diverse community of scholars. Our charge was to:

- understand AI research and applications already underway across the university;
- develop a vision for how Yale can lead and innovate in AI research and influence its positive applications for society; and
- recommend coordinated action to amplify the AI-directed efforts of each Yale school.

From January to March 2024, our Task Force met with each of the dean- and unit-led panels, gathering input from across schools, disciplines, and administrative units. Through this process, we conferred with more than one hundred faculty, staff, and students who are developing, engaging with, and considering the impacts of AI in disciplines and areas of operation that span the university.

The following pages outline thematic insights that resulted from our conversations with these Yale colleagues and from our own deliberations as a Task Force. Based on our interactions, in this report we reflect on ways that Yale is strategically positioned to lead AI development, and we recommend areas for investment and coordinated, pan-university action, with the goal of providing a framework through which Yale can lead.

Themes and Observations from University AI Panels

Panel presentations from various Yale schools and units demonstrated rich disciplinary distinctions among their approaches to AI. Nevertheless, several cross-cutting themes and observations emerged as foundational to Yale's vision for developing and addressing this technology. These values, ideas, and principles, which recurred throughout our conversations, inform the recommendations found later in this document.

Yale is home to many existing efforts related to AI.

In many disciplines throughout Yale, extensive and varied work with AI is underway. Yale faculty, postdocs, students, and staff are advancing groundbreaking research related to AI in many fields.

Yale can actively develop and examine transformative implications of AI.

Rather than passively wait for AI to develop, Yale can continue to advance the core discipline of AI, draw on Yale's strengths across disciplines to drive application-driven AI research, contribute to interactive AI, and lead in understanding and shaping AI development. Yale can provide tools and infrastructure to enable productive application of AI to all disciplines.

Students must be prepared to lead in an AI-infused future.

Robust experimentation with AI in teaching and learning is taking place all over campus. Yale needs to evolve its curricula to prepare graduates for a society and workforce where AI skills are becoming essential, and to provide consistent and clear criteria for the appropriate use of AI in teaching and assessment. The Poorvu Center for Teaching and Learning serves as a nexus for faculty and student support.

Yale is positioned to lead in the exploration of ethical and social impacts of AI.

The technology industry will prioritize AI products and features that it can monetize. Guided by its mission and values, Yale can be a leader in advancing the understanding of the ethical and societal impact of AI as well as in developing new approaches to promote equity in or through AI technologies.

Cross-cutting catalysts will fuel AI development.

Although many implications of AI vary by discipline, its development also transcends traditional academic boundaries, structures, and funding models. Investment in initiatives that span the university will help foster collaboration across fields.

Yale's approach to AI must account for generational differences.

AI is developing rapidly across many fields, and we have found disciplinary and generational disparities among the ways that faculty utilize AI in their research and teaching. Members of the Yale community will need support and encouragement to utilize and develop AI skills effectively, creatively, and safely. Broad-based support, tailored to the different needs of our diverse faculty members, will provide important opportunities for everyone to develop greater familiarity with the possibilities presented by AI.

Yale must responsibly manage its institutional data and take advantage of the insights AI can yield, particularly given the university's access to quality data.

Access to quality data at scale is necessary for advancing AI. Yale has a notable advantage with high fidelity proprietary data sets from, for example, its vast museum and library collections as well as clinical data from a diverse patient set served by Yale Medicine and Yale New Haven Health. In addition, new initiatives such as the Data-Intensive Social Science Center and the Yale Center for Geospatial Solutions are focused on constructing valuable data infrastructures. This advantage comes with a responsibility to protect privacy and model the responsible use of data. In many areas of the university, work is needed to improve the consistency of policies and practices for managing institutional data, to prioritize infrastructure that transcends fields, and to support the acquisition of large datasets in order to realize the opportunities AI presents.

Yale's Unique Strengths

Beyond these general themes, the Task Force was inspired by the panel presentations, which highlighted a unique collection of strengths that Yale can bring to bear on the worldwide development of AI. Yale's faculty bring expertise to AI from a wide range of disciplines, not just from computer science, engineering, informatics, and data science. The university's historic excellence in the arts, humanities, and social sciences also allows us to situate current AI developments in a deeper historical context and explore its affective, creative, and human dimensions in novel and unique ways.

Our ability to shape the development of AI is greatly enhanced by university investments, many of which predate the wide public availability of ChatGPT and other similar generative AI tools. For example, in July 2022, Yale developed a distinct faculty for the School of Engineering & Applied Science, leading to new faculty lines in computer science and other fields, which have attracted a number of scholars working on foundations and applications of AI. Slots for additional hires in these areas are available, and joint appointments with other schools and departments are encouraged to attract and empower scholars interested in deep interdisciplinary work related to AI. This is also true in the Faculty of Arts and Sciences' Department of Statistics and Data Science, the School of Medicine's recently created Department of Biomedical Informatics & Data Science, and the Wu Tsai Institute, where new compute clusters and faculty lines that emphasize AI have been created. Through other recent investments and initiatives, including the Digital Ethics Center, Project Lux, and the Yale Center for Geospatial Solutions, the Yale community is well positioned to promote understanding of the human dimension of AI.

Yale can leverage these strengths to develop AI technology, refine models so that they adapt to fields, make computation more efficient, and correct deficiencies or bias. In addition, we believe that Yale can examine AI's societal impact and counteract potential negative aspects of AI's advancement, such as environmental harm or threats to democratic values. These are just a few of the ways that Yale can be a leader in the advancement of AI technology and its ethical development.

Recommendations for Investments and Coordinated Action

Yale's schools are the primary locus of AI development and activity for their disciplines, and the dean-led faculty-panel summaries highlight the breadth of activity already underway. Yale's AI strategic plan is centered at the school level, but we also observed themes extending beyond schools and units, themes for which central investments are needed to achieve the vision we seek. To amplify school and unit efforts, we recommend a set of coordinated investments in AI infrastructure. These investments reflect the themes we discussed with our colleagues across campus and are intended to build upon Yale's unique strengths as a university.

As commercial entities, technology companies are investing substantial sums in artificial intelligence to develop commercial products. We recognize that Yale cannot, and should not, take the same approach. Rather, our AI efforts should be guided by Yale's mission to improve the world today and for future generations. To do so, we need to scale and bolster in-house computational

capacity, AI expertise, and access to generative AI tools that enable our community to experiment, learn, and positively impact the development of AI. Because AI technology is advancing rapidly, we suggest that investments be flexible, allowing for responsive adjustments over time. We make the following recommendations for coordinated investment and action.

Compute

Yale should invest in a cluster of at least a few hundred advanced graphics processing unit (GPU) to enable faculty to advance research on AI models from within Yale. Relative to the need to seek computational resources externally, having this capacity at Yale would provide faculty with greater priority, control, and speed to conduct research. Special consideration should be made to accommodate processing of high-risk confidential data, such as those underpinning medical research and discovery. We recommend phasing this investment over several years given that faculty needs will grow, supply of GPUs is limited, and GPU technology is advancing rapidly.

In addition to hosting dedicated GPUs at the university, Yale should purchase cloud-GPU hours from technology providers to supplement its capacity. Though more expensive per unit, these hours can be scaled up or down to respond flexibly to faculty needs. Of course, these investments will need to be made in the context of the university's climate goals, presenting another opportunity for Yale to lead in environmentally responsible AI. Yale should also explore opportunities with industrial partners for compute capability and access to technological innovations.

For the community to make maximal use of this computational investment, Yale should create new AI-focused research-scientist positions, possibly through the Yale Center for Research Computing, to serve as a core resource. The scientists who fill these roles should partner with faculty and trainees to advance their research agendas.

Access to generative AI tool(s)

To encourage experimentation and learning, Yale should begin by providing equitable access to a leading, multimodal generative AI tool to all members of the Yale community. We recommend that Yale procure or develop access to a generative AI tool that will keep individual and university data private. This is such a rapidly changing space that attention must be given to flexibility over even short periods of time. The university might consider making more than one tool available to Yale faculty, staff, and students or offering flexibility in what tools are made available for use.

Training

Yale should develop or procure diverse, targeted trainings for faculty, staff, and students. These trainings should support individuals in developing basic, intermediate, and advanced skills to apply AI in their research, teaching, learning, and work. They should take multiple formats and be tailored to different types of disciplinary needs, making them as accessible and effective as possible. Formats may include faculty workshops, prompt-a-thons, on-demand virtual courses, and master classes.

Catalysts for collaboration

To encourage cross-disciplinary partnerships, Yale should convene new events related to AI and consider associated incentives for participation. For example, Yale research symposia on AI, with panels and poster sessions, could highlight discoveries that are underway and encourage new partnerships. Such symposia or seminar series could be accompanied or supplemented by seed grants offered to faculty from different departments partnering on a research project. When appropriate, these initiatives should leverage the range of expertise at Yale, fostering interdisciplinary collaborations.

Curricular innovation

To attract the best students and prepare them for an AI-infused future, faculty, departments, and professional schools should continue to refine or develop new courses, majors, or degree programs related to AI. As an educational institution, Yale has the opportunity to serve as a model by adapting its curricula to prepare students to lead in advancing and considering the implications of AI.

Policy

Yale should refine its policies to clarify the proper use of AI with respect to teaching and learning, research and publishing, university operations, and institutional data and software/tools. The policies should encourage experimentation and learning while safeguarding privacy, security, and proprietary assets.

AI for humanity

Yale's excellence in the humanities and arts — ranging from philosophy to linguistics to drama — should be leveraged to guide AI development, practices, and policies that are not only ethical but beneficial to society. While industry prioritizes technology for consumption and profit, Yale can at once advance AI innovation and develop the social, legal, and ethical frameworks that guide its use. As an institution known for strong interdisciplinary collaboration, Yale should consider robust incentives for scientists and engineers to partner with humanists and artists to develop and address AI that considers and confronts, rather than circumvents, the concerns it raises.

The university should also ensure that its students — the next generation of scholars, industry leaders, and citizens — are equipped with a wide range of skills needed to contend with technology. As these individuals venture into the world post-Yale, they should be uniquely prepared to weigh the benefits of AI with the issues it may present to humanity. Yale's schools might consider developing lists of core competencies that every graduate should have upon completion of their degree.

Conclusion

AI represents a generational advancement in technology with wide-ranging implications for all aspects of Yale's mission, including its intellectual endeavors and its operations. Rather than waiting to see how AI will develop, we encourage our colleagues across Yale to proactively lead the

development of AI by utilizing, critiquing, and examining the technology. We see vast potential implications of AI for Yale and humanity.

In the course of our work as a Task Force, we have heard about the significant, distributed AI efforts by Yale colleagues across the university. We are encouraged by the efforts the deans are leading in their schools, as highlighted in their written summaries. To amplify these efforts, we recommend university-wide investments, opportunities for cross-disciplinary collaboration, curricular innovation, and policy development. We recommend that, during the summer of 2024, the university develop a detailed plan for institutional investment that can allow the implementation of these priorities. With these investments, and by drawing on the community's collective expertise and ingenuity, Yale will lead in shaping the development of AI for the good of the university and, more broadly, for humanity.

Acknowledgments

The Task Force is grateful to the many colleagues at the university who supported its work. In particular, we thank the dean- and unit-led panels for sharing their ideas and insights with us. We also thank Task Force members Jenny Frederick and Chad Losee for creating initial drafts of this report. Finally, we are grateful to Vin Guerrero and Mehmed Can Olgac for staffing the committee and to Lindsay Guarino for managing the committee's scheduling needs.

The following summaries were authored by individual schools and units. The summaries address school-, unit-, and field-specific visions for advancing AI as a technology and a subject of scholarly inquiry; using AI tools in research, education, and operations; and shaping the social and ethical impacts of AI for the good of humanity.

Appendices

Appendix 1: AI Summary—Education

Appendix 2: AI Summary—Yale School of Architecture

Appendix 3: AI Summary—Yale School of Art

Appendix 4: AI Summary—Yale Divinity School

Appendix 5: AI Summary—David Geffen School of Drama at Yale/Yale Repertory Theatre

Appendix 6: AI Summary—Yale School of Engineering & Applied Science

Appendix 7: AI Summary—Yale School of the Environment

Appendix 8: AI Summary—Faculty of Arts and Sciences: Humanities*

Appendix 9: AI Summary—Faculty of Arts and Sciences: Basic Sciences*

Appendix 10: AI Summary—Faculty of Arts and Sciences: Social Science*

Appendix 11: AI Summary—Yale Jackson School of Global Affairs

Appendix 12: AI Summary—Yale Law School

Appendix 13: AI Summary—Yale School of Management

Appendix 14: AI Summary—Yale School of Medicine

Appendix 15: AI Summary—Yale School of Music

Appendix 16: AI Summary—Yale School of Nursing

Appendix 17: AI Summary—Yale School of Public Health

Appendix 18: AI Summary—Yale College

Appendix 19: AI Summary— Collections and Scholarly Communication

Appendix 20: AI Summary— Clinical Practice

Appendix 21: AI Summary— Operations

*The Graduate School of Arts and Sciences is represented in the plans of the Faculty of Arts and Sciences.

AI in Education at Yale

Summary of Themes and Recommendations by and for the Yale Community

Jenny Frederick, Associate Provost for Academic Initiatives

MARCH 2024

In December 2022, the Poorvu Center for Teaching and Learning began fielding inquiries from faculty about ChatGPT and teaching. By January 2023, we had created a website with guidance and resources and begun hosting events for faculty to discuss generative artificial intelligence (AI) in education and exchange ideas. More than a year later, our AI website is in its sixth edition. Many instructional consultations feature ideas and questions about AI tools. Our peer-tutor training has evolved to address appropriate guidance about AI use in assignments. This evidence demonstrates how rapidly AI technologies have swept into education at Yale and the widespread need for reflecting on AI's influence on our educational mission.

In spring 2024, the status of AI in education at Yale varies by school, discipline, and course. Many of the deans' AI research panels addressed elements of AI education, highlighting the need for AI literacy and field-specific applications. Although numerous instructors are integrating AI into their courses, some have yet to do so. New courses on AI are offered in the Schools of Medicine, Architecture, Divinity, and Yale College, to name a few. Students across Yale recognize that AI knowledge and skills will be essential when they graduate and enter the workforce. They take to heart the idea that AI will not take jobs, but humans who learn AI will.[1] However, we have yet to significantly rethink curricula for Yale's many schools and programs. The central question Yale University must address is what our students need to learn in a world with AI. Our responses to this question can ensure that we maintain our strong reputation for educational excellence.

This summary was informed by well over a year of conversations with faculty groups and committees, with individual instructors and staff colleagues all over campus. Below is a list of groups consulted:

- Academic affairs deans from schools and colleges
- Poorvu Center Faculty Advisory Board
- Poorvu Center Student Advisory Board
- Faculty Committee on Digital Education

1 This provocative idea is attributed to Kai-Fu Lee, a prominent computer scientist and AI expert who wrote a 2019 book titled *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin, 2018). After generative AI became widely accessible in late 2022, many public AI figures have restated this idea in articles and interviews.

- Yale College Committee on Majors
- FAS/SEAS Faculty Senate
- Yale University Library
- Tsai CITY
- Many Yale leaders, faculty, and students

The first two groups listed above were formally charged by Provost Strobel to address the following questions about AI in education:[2]

1. What are the opportunities and challenges of integrating AI into our educational mission?
2. Which pilots should Yale run to learn more? What support do you need for these pilots, such as funding, staff time, education and training, access to tools, etc.?
3. What adjustments to policy or guidelines would support these efforts?

Despite the breadth and variation of education at Yale, common themes emerged from these discussions. Below are summarized responses to each of the questions, followed by key themes and accompanying recommendations for Yale's approach to AI in education.

1. Challenges and Opportunities

AI raises challenging issues of academic integrity. Early reactions to AI in higher education were characterized by concern that students would circumvent their learning by relying on AI tools to complete assignments. While that concern has been largely unrealized, rapidly improving AI output means that instructors need to develop and justify AI use policies for their students. Because AI is a predictive software, responses to prompts often include false information known as "hallucinations." Flaws in training data also produce responses that can be biased and outright wrong. While these limitations create learning opportunities, they can also be problematic when not handled appropriately. Another challenge arises from AI's constant evolution, and new tools, features, and industry developments make it hard to keep up. It takes time to gain and maintain AI literacy, and faculty engagement is uneven. Overall, students tend to be ahead of faculty in their facility with AI use. They are receptive to guidance, although they may experience confusion when AI is handled differently from one course to the next. Because AI tools and issues vary by discipline (writing versus coding; medical education), rethinking education requires careful, field-specific examination of courses and programs. Students will need to have AI knowledge and skills when they leave Yale, but our curricula do not yet reflect that expectation. Finally, most conversations about AI recognize several divides that affect Yale's ability to provide a world-class education. There are those between faculty who integrate AI tools or topics into their teaching or research and others who do not. There are divides between students who pay for premium subscriptions and those who cannot afford it. The lack of equitable access to high-performing AI presents a barrier to students as well as faculty.

2 On November 28, 2023, Provost Strobel wrote to all deans and leaders in education, cultural heritage, and practice (clinical and operations). He asked schools and relevant groups to discuss the three questions listed in this narrative and share their reflections and recommendations with the Yale Task Force on AI.

The challenges described above contrast with exciting educational opportunities related to AI. There is a clear need to emphasize liberal education as a source of understanding and habits of mind to navigate the world with AI. We can shape education with a renewed emphasis on deeply human skills, such as critical thinking, asking questions, drawing from historical and philosophical traditions, applying ethics, and being creative. Yale has a world-class faculty, and we can lead the way to demonstrating how institutions can adapt. Faculty expertise is needed more than ever since domain-specific knowledge is essential for evaluating AI's usefulness. AI offers a vast trove of knowledge, but faculty leaders are essential to help us harness the power of AI for the good of humanity and to demonstrate for students/trainees how best to apply it or avoid it in their fields of interest. Faculty and students can work collaboratively to explore AI opportunities and limitations. Yale can also ensure that its approach to AI reflects the university's diversity, equity, inclusion, and belonging (DEIB) values. That means we need to consider questions of access, how investments reflect values, whose voices are involved in decision making, who is represented in AI leadership, and what opportunities are available to all students and faculty. AI tools to augment courses and learning have the potential to level the playing field for students. Individualized learning-support will allow students to complement work with peer tutors or prepare for class or office hours. We have the opportunity, and responsibility, to reflect deeply on discipline-specific education and ask, "Where does the thinking happen?"[3] How can we preserve essential elements of learning in a discipline and consider how students might enhance their learning with AI tools? The world may look to Yale as a leader in articulating educational priorities to prepare students for a world with AI.

2. Pilots and Support

Instructors in many schools and departments began exploring AI in their courses in spring 2023. Examples span the disciplines, such as the following creative approaches:

- Comparing AI-generated writing to human writing and discussing the craft of writing;
- Coding with AI tools so students can focus on more nuanced details;
- Working on a problem set with AI assistance, then working with a small group in class to correct inaccuracies and refine answers;
- Assigning AI collaboration in order to identify counterarguments to strengthen a persuasive essay;
- Starting with AI-generated text and documenting revisions to reveal the writing process.

Instructors with more AI experience have assigned projects where students apply knowledge of large-language models (LLMs) to augment their entrepreneurial projects. A language instructor was awarded a teaching grant to develop AI-based, interactive grammar lessons. In areas such as architecture and medicine, industry expectations have shifted, and new courses are being launched to meet them. In contrast to these examples, some faculty have not engaged with AI and have policies that prevent their use.

3 This question was posed by Johann Neem in an October 11, 2023 *Inside Higher Education* essay by the same name. https://www.insidehighered.com/opinion/career-advice/teaching/2023/10/11/academe-should-make-discipline-specific-responses-chatgpt

Not surprisingly, many students are engaging with ChatGPT and other AI tools. Students are eager for guidance from faculty and want to use AI responsibly. Many are discovering creative ways to support their learning with AI, including the following examples:

- Asking ChatGPT to summarize a topic in order to check their understanding;
- Asking for alternative explanations when a lecture or concept is unclear;
- Using chatbots to quiz themselves and provide coaching as they learn a topic or practice a skill;
- Studying foreign language vocabulary with assistance from a chatbot.

In addition to routine learning uses, some students with disabilities benefit from AI capabilities. There are many ways to use AI tools to enhance accessibility and support executive function, such as for those enlisted by students diagnosed with ADHD. Students learning English as a second language or who are unfamiliar with norms of academic expression are relying on chatbots to check their work or help generate an appropriately worded email to a professor. These tools help more students gain confidence and thrive in their studies at Yale.

Another type of pilot is a "walled garden" chatbot set up for several spring 2024 courses. These pilots provide AI tools that are managed privately within Yale's security perimeter, rather than relying upon the open commercial product, where inputs are used to train the model. Instructors have varied reasons for requesting a walled garden pilot. In some cases, their students will be using nonpublic data. In other cases the course topic may be sensitive, and the instructor prefers to create an environment where there is no threat of corporate intrusion or capturing of student prompts. These walled gardens are currently supported by a collaboration between Information Technology Services (ITS) and the Poorvu Center and require nontrivial technical work to set up. We have more requests than we can accommodate at this time, and lessons from the pilot will inform ongoing efforts. Schools and departments are beginning to discuss curricular adaptations that may be needed to adequately prepare students in AI topics. We need to do this comprehensively so that all students are well equipped with general and discipline-specific AI knowledge.

3. Consideration of Enabling Policies and Guidelines

Showcasing examples of AI integration in courses and curricula is one way to inspire faculty to consider how they might approach AI in their courses. Faculty are also engaging in virtual and in-person forums to exchange ideas and troubleshoot with each other. The Poorvu Center's website features examples from Yale faculty and related resources. Yale instructors are free to determine the extent to which AI can be used in their courses, within usage guidelines. Faculty are encouraged to communicate AI use policies to students in their courses. Many have found the following three-part framework from Cornell's 2023 report on AI a helpful starting point (see Figure 1).[4] An instructor

4 The framework in Figure 1 was adapted from the "CU Committee Report: Generative Artificial Intelligence for Education and Pedagogy," spring 2023. https://teaching.cornell.edu/generative-artificial-intelligence/cu-committee-report-generative-artificial-intelligence-education

may ask students not to use AI in cases where they need to practice and develop a particular skill. Long division illustrates this point, since it is rarely used but consistently taught in primary schools because it teaches foundational concepts of factors and divisibility. Instructors may also give students permission to use AI with the expectation that they document this collaboration. This approach can help instructors learn from their students and promote intergenerational partnerships within a course. In other cases, instructors incorporate AI tools in specific assignments designed to advance learning in that course.

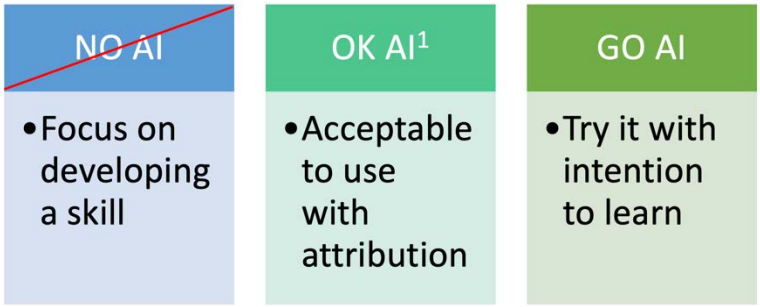


Figure 1. A framework instructors may consider as they develop course policies for AI use.

To enable curricular revision, we need to identify incentives and support that will enable schools and departments to take on this work. Expectations from Yale leadership can go a long way in signaling permission for such efforts. In the Cultural Heritage panel, University Librarian Barbara Rockenbach suggested a "service layer" approach to coordinating how-to resources, guidance on tools, and triaging teaching and research inquiries. To support faculty and students, we need to provide curated educational materials on AI literacy. The library and the Poorvu Center could partner to launch such a project. In addition to identifying excellent existing materials, developing Yale-specific AI literacy resources would help meet the needs of our teaching and learning community.[5] Finally, we need to provide licenses that enable all students and faculty to access best-in-class AI tools relevant to their disciplines.

Summary of Themes and Recommendations for AI in Education at Yale

1. The need to democratize access to best-in-class AI tools

    A. Invest in licenses to AI software to enable all faculty, students, and staff the ability to learn and experiment.

    B. Maintain flexibility in our AI software commitments, since preferred tools will vary by discipline and new front runners will emerge.

5 See, for example, "Creative and critical engagement with AI in education," a guide created for Harvard's AI Pedagogy Project; and YouTube video tutorials by Lilach Mollick and Ethan Mollick, such as Practical AI for Instructors and Students Part 1.

2. The need for education about AI.

    A. Provide curated resources in a variety of formats to promote AI literacy.

    B. Address AI in our curricula to prepare students for ethical leadership and service in the world with AI.

3. The need to center our DEIB values, to counteract bias, exclusion, and other harms that arise from how AI systems are created and trained

    A. Integrate ethics of AI use into our educational approach.

    B. Promote interdisciplinary initiatives that provide multiple lenses to consider the role AI should play in the future of humanity.

Yale ARCHITECTURE

Report to the Yale Task Force on Artificial Intelligence

MARCH 2024

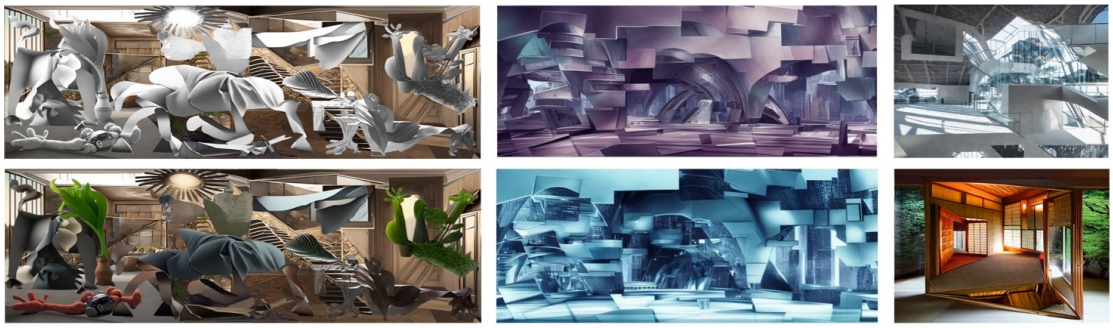
In a recent article in *The Atlantic,* Washington University computer scientist Ian Bogost asks, "whether computing ought to be seen as a superfield that lords over all others, or just a servant of other domains, subordinated to their interests and control. This is, by no happenstance, also the basic question about computing in our society writ large."[1] Computing has had a profound effect on the design and construction of the built environment from the earliest advent of computer-aided drafting through to today's digital fabrication methods, and tomorrow's autonomous algorithms will continue this trend. But architects are not mere consumers of new kinds of computing but rather important innovators. Much of the fundamental thinking that has brought the modern age of machine learning has strongly architectural roots:[2] Christopher Alexander's *Pattern Language* established principles of object-oriented programming;[3] Nicolas Negroponte's MIT research in computational design anticipated today's intelligent systems;[4] and Richard Saul Wurman defined early versions of today's "big data" with his work in information architecture.[5]

Architects create our spaces of human habitation through design as a discipline that fits squarely between traditional humanities and arts (like music, art, or drama) and the sciences (engineering, environment, public health, medicine). Architects further navigate between considerations of law and policy, material science, and supply-chain and labor economics (to name a few), and spatial design demands an understanding and integration of many perspectives. A clear strategy will be required for AI to properly serve the fundamentally multidisciplinary nature of architectural practice that must include strong reliance on the work of other researchers and practitioners. That strategy should include, in the words of composer Katie Balch, convergent and divergent components, to support both the analytical and generative demands of design: analysis that supports the quantitative inputs to design and generation of aspects of design itself.[6]

1 Ian Bogost, "Universities Have a Computer-Science Problem: The Case for Teaching Coders to Speak French." *The Atlantic* (March 19, 2024). https://www.theatlantic.com/technology/archive/2024/03/computing-college-cs-majors/677792/.
2 For a superb summary of this history, see Molly Wright Steenson, *Architectural Intelligence: How Designers and Architects Created the Digital Landscape* (Cambridge, Mass.: MIT Press, 2017), xii.
3 Christopher Alexander, Sara Ishikawa, and Murray Silverstein, *A Pattern Language: Towns, Buildings, Construction* (New York: Oxford University Press, 1977), xliv.
4 Nicholas Negroponte, *The Architecture Machine: Toward a More Human Environment* (Cambridge, Mass.: MIT Press, 1970).
5 See Steenson, *Architectural Intelligence,* chapter 3, "Information, Mapping, and Understanding."
6 Professor Katie Balch of the School of Music, Yale AI Task Force meeting, February 2, 2024.

The School of Architecture is well-positioned to develop such a strategy and test its characteristics. We are a global leader in design education, attracting the world's best students, who are instructed by practitioners and academics who are the leaders of our profession. We support these efforts with what is largely considered to be the best technological infrastructure — computing and machinery — in design education today. We work in digital image generation, computer-controlled fabrication, and robotic assembly routinely, and our students are equipped with a broad array of the latest software at every studio workstation as they move fluidly among tools and technologies. These machines are sufficiently powerful to accommodate the next generation of AI-based software.

Our current efforts represent the early days of the development of AI in architecture, and the implementation of such technologies in the profession at large are also nascent. We have created two internal generative platforms for experimental use by our students: Projects Minerva and Vulcan. The former is the basis of an internal technology help desk that has helped us understand the basis of text-based generators; the latter is a research platform on which we are currently testing the ability of text-based systems to read and understand building codes and take professional certification tests. We have made a generative tool, based on Comfy UI, available to students to experiment with image generation, and many are carefully testing such systems for their studio design projects. The lack of input and output control in such systems in their current incarnation makes their utility limited, but students are investigating them daily in our studios.



Images from Altered States of Architecture Seminar
YSOA Spring, 2023

We are teaching courses that directly address questions of machine intelligence in the profession. Senior Critic Brennan Buck has taught three successive classes in past years on topics of machine autonomy and design output, including AI Aesthetics (2021), Altered States of Architecture (2023), and The Black Box: Architecture in the Age of Opacity (2024). Professor Adjunct Phil Bernstein and Lecturer Sam Omans are currently teaching the second iteration of a course on the implications of AI for the practice of architecture, Architecture and Machine Intelligence in Theory and Practice (2023, 2024), based in part on Bernstein's 2022 book.[7] These instructors, working with Senior Director of Advanced Technology Vin Guerrero, are currently developing a larger-scale course for the next academic year designed to give students a foundation in the principles of AI computing, set

7 Phillip G. Bernstein, *Machine Design: Architecture in the Era of Machine Intelligence* (London: Royal Institute of British Architects Publications, 2022).

the historical and theoretical basis of its use, examine the current developments in the commercial marketplace, and provide opportunities for experimentation and research using AI. The Poorvu Center is assisting with course development. Finally, Vin is a current thought leader on campus in AI technologies, particularly on behalf of Yale's Information Technology team.

General purpose AI offerings, while useful for such tasks as text generation and editing, are currently not of great utility in the specialized, multidimensional work of the architect, and domain-specific platforms and tools must emerge to truly serve the profession and the building industry at large. It is likely that it will be some time before such tools are readily available, and even intermediate "building-specific" platforms on which such tools can be created are yet to be developed. In their place, there are currently almost 1,000 putative "AI-based" software tools in the building industry marketplace suggesting that the early days of AI development in architecture will be characterized by market-driven chaos, making it even more important to establish principles and directions from which architects can take advantage of current developments and guide future ones.[8]

If today's generative technologies suggest a direction for Balch's "divergence," the "convergent" tools for architects, engineers, and builders largely do not exist today. The analytical inputs to spatial design are diverse, ranging from structural stability through issues of project costs and environment impact, to name a few. The data from which such tools can be developed is diffuse and disorganized, and principles of data coherence have yet to be established in our industry, which has much to learn from other disciplines—including many that the Task Force has seen during its work—about how to marshal and leverage information toward analytical ends.

Thus, there are great opportunities for Yale to establish the most important theoretical and practical principles on which architects can and will use AI appropriately. We believe we can do so with our current resources through teaching, research, and professional leadership, augmented by support from the University. Architecture will never operate at the level of scale or sophistication of law, medicine, or public health in terms of research projects and consumption of computation resources, but we do believe we can provide leadership precisely because of our pivotal role at the intersection of the arts and sciences. As the campus develops its AI approach, that support might include the following:

- Structured access to compute resources that might otherwise be consumed by larger enterprises on campus. It will be critical to make such resources equitably accessible by smaller units like ours.
- A central campus center, similar in structure and approach to the recently founded Geospatial Solutions Center, at which we can learn best practices, follow what is happening elsewhere on campus, receive training, and access necessary expertise.
- As has been discussed, provide access for our students and faculty to general purpose generative tools. As 90 percent of our students are supported by financial aid, and most of our faculty are

8 See "AI in AEC | AEC AI Hub" at https://aecaihub.addpotion.com/

practicing professionals, such access will level our pedagogical playing field across the school for AI work.

- Support for designed, structured relationships with key collaborators across campus — particularly in engineering and environment, where projects of common interest in spatial design, material science, and urban planning can be explored, possibly through funded research projects.

Finally, the school has limited resources for hiring full-time faculty, and it is at a particular disadvantage in hiring a colleague who could work across AI computation and architecture given the intense competition for such academics and the salary structures in architectural education. We believe one of the greatest opportunities to accelerate our work in AI would be a cross-appointed faculty member, likely in computer science and possibly working in Julie Dorsey's group, and architecture. This would match our current strength in technology theory and professional practice with computational expertise and allow us to begin setting the pace for architectural pedagogy, AI, and the profession at large — exactly what Yale should be doing for the profession and discipline of architecture.

Yale School of Art

Report to the Yale Task Force on AI

By Anahita Vossoughi, Alvin Ashiatey, Benjamin Donaldson

APRIL 1, 2024

The School of Art is committed to fostering exploration across a diverse spectrum of tools while embracing an experimental, critical, and multidisciplinary approach. The emergence of generative artificial intelligence has sparked discussions about the essence of art and the implications and benefits of AI for artists. The school is in a unique position to introduce students to these new tools. Demystifying the complexities of generative AI, including its mechanisms, limitations, and failures, can help students gain a comprehensive understanding of how predictive machine-learning algorithms can function within the creative process. This educational approach not only cultivates technical proficiency but also encourages an emergent, thorough, and critical examination of the role of technology and, specifically, AI in art.

Current Uses in Research and Teaching

Currently, the School of Art offers a couple of courses that incorporate AI into its curriculum. In the graphic design department, a faculty member experimented with teaching image-generative AI tools such as Stability AI's open-source Stable Diffusion models. Meanwhile, another faculty member in the painting department focused on leveraging Adobe's AI tools, such as Adobe Firefly, to expand ideas surrounding painting, collage, and appropriation.

Through the Interdepartmental Days programming and the graphic design department, the School of Art has also presented to the community thought leaders and artists who utilize AI in their practice. Studio visits and critiques from artists who use AI in their practice have been made available to students and are expected to expand and continue.

Additionally, the School of Art is implementing its multiyear lab restructuring, which will introduce AI and chatbots into the school's general digital media support and training offerings in FY25.

AI Accessibility and Opportunities at the School of Art

Our faculty see a few pathways to provide access to training in using AI as part of the dean's vision for the Ethical MFA program: the hardware (the equipment and software), the workforce (the people), and the program (the education).

The Hardware: AI computing environment in the All-School Making Lab

The All-School Making Lab at the School of Art is projected to be installed in FY25 as part of the school's transition from analog equipment to digital tools. To enable students to delve beyond generic AI resources, which often can be fraught with problematic images sourced from the internet or constrictive guardrails, we aim to transition to a model where we can offer a robust AI computing environment supported by staff and faculty that is available for classes and students. This setup would empower us to curate our own datasets and experiment with machine learning with oversight from AI practitioners. This would enable our community to explore with specificity and expand the limits of what is achievable with open-source generative AI.

In terms of a recommendation for university-accessible hardware, an independent server indexed and isolated from the rest of the internet community with reliable access, would allow for a more precise use of these tools and give the community a better sense of control and authorship based in the discipline of each field. Imagine faculty and students ranging from the School of Art to the School of Medicine to the School of Management contributing to an independent reference pool to which the Yale community (and its research partners) have access. It shares similarities to the transferrable model of peer reviewing and journals/publications databases that drive the research of higher education institutions; content is driven by the discipline with standards to maintain academic rigor and integrity.

The Workforce: AI practitioners

The school would seek to increase its human resources, particularly looking for dedicated talent proficient in coding, AI, and computer maintenance to ensure the smooth operation and support of an AI lab and teaching environment. This might be a mix of staff, faculty, and artists in residence who also contribute to other needs across the school. The faculty and staff who currently work with AI would also be enabled to expand their own proficiency in AI while being supported with new human resources. Access to a central set of resources and expert technicians would benefit the school as well.

The Program: Educating the community

The way AI draws from and uses text and images raises many copyright and authorship questions. The school could partially respond to those concerns with specific authorship, best done by drawing from curated, user-made images or text pools. As the societal conversation around these questions continues to evolve, the guidelines that Yale follows should allow for experimentation by artists through access to greater controls by the users.1 The introduction of a series of workshops focused on technology, designed specifically for artists with little or no experience in this area, would serve as a bridge connecting artists to the rapidly evolving field of AI. We see great value in inviting critical thinkers from the nexus of art and technology to engage with our students through

1 Zachary Small. "As fight over A.I. artwork unfolds, judge rejects copyright claim." *New York Time*s (August 21, 2023). https://www.nytimes.com/2023/08/21/arts/design/copyright-ai-artwork.html

workshops or talks. This would help raise awareness of ongoing art world dialogues surrounding artificial intelligence and expose students to diverse perspectives and discourses in the field. Examples of such thought leaders are James Bridle, Boris Groys, Ramon Amaro, Kate Crawford, Ruha Benjamin, Alexander R. Galloway, and Lev Manovich.

As a final point, central support for training around the ethical and legal implications of using AI in cutting edge research would also serve the school. Students and faculty will have questions that evolve alongside the technology as it develops.

Quotes from MFA students

> (1) In my work, I primarily involve custom Machine Learning, which varies from project to project. To better explain, I'll describe a specific project titled "Command and Freedom." In this project, I created a custom AI Language model based on a collection of my personal data over three years. This data includes all my emails, messages, books I've read, movie transcripts, advertisements, and administrative documents, etc. . . . The system asked unanswerable philosophical questions, such as "Have you ever disappointed the person you loved?"—highlighting the complexities of self-reflection and personal data interpretation. I was interested in exploring the relationship between self-reflection, our thoughts (including virtual ones), data collection, and the limits of Language models.
>
> (2) Unfortunately, in the School of Art, not many students are deeply engaging with AI. I believe this is due to the steep learning curve associated with training custom models, among other challenges.
>
> (3) I've been more interested in organizing and creating frameworks for people to discuss AI and the issues surrounding it—almost more as an editor/facilitator than an artist—and personally, at the current stage of this tech, I am way more excited and enthusiastic about talking to other people about how they are working with it than working with it myself.

The Bottom Line

The School of Art community is already embracing this technology by its own means on the individual level; more centralized and dedicated resources to making AI a resource at the forefront of the MFA education will strengthen the school's capacity to address this paradigm-shifting technology and to equip artists who seek to utilize it as leaders in AI.

AI at Yale Divinity School

APRIL 1, 2024

Plato reported that the pre-Socratic philosopher Heraclitus said that a person "could not step twice into the same river" (*Crat.* 401d–402c). AI has accelerated the speed of the river so that one feels as though the river changes even as a person's foot is going into it.

As a professional school based in the humanities and the arts, we have used AI unevenly in research, teaching, and administration. Beginning in 2023–2024, however, we have made efforts to explore AI formally at Yale Divinity School (YDS) in five ways. First, we have discussed the use of large-language model chatbots (LLMs), such as ChatGPT, in course instruction two times at faculty meetings. We decided to follow the principle of subsidiarity and allow each instructor to set guidelines for their own courses but requested that they include these guidelines in their syllabi. Second, we held a special retreat devoted to AI in January 2024 with the leadership of the school and select faculty. Every attendee was expected to have read Mustafa Suleyman's *The Coming Wave: Technology, Power, and the Twenty-first Century's Greatest Dilemma* (2023). We invited Jenny Frederick to lead most of the afternoon retreat, which went so well that we invited Dr. Frederick to make a presentation at our February staff meeting and provide an updated overview of the university's initiatives on AI. It was well received. Third, we are devoting the spring issue of our online magazine *Reflections* to the topic of AI. We are asking faculty, students, alumni, and others from Yale to address some of the concerns that we will signal below. Fourth, we experimented with AI in the second chapel service of this semester. Associate Dean Awet Anademichael created a service using ChatGPT and read a homily created via ChatGPT in her style. She solicited feedback from the audience afterward. The consensus was that ChatGPT did a respectable job of capturing her vocabulary, but offered a flat sermon, filled with platitudes but lacking substance and surprise (almost predictable at this stage of the development of ChatGPT). Finally, we are in the process of scheduling a ninety-minute training session with AI for staff in concert with Frank Matthew (IT) and John Baldwin (Library), who will build a "walled garden" that will permit us to experiment with administrative possibilities without exposing any sensitive data beyond Yale. We will invite twenty staff from different offices across the school to this event.

What could a divinity school contribute to the university's response to AI? In all candor, at this point we have more questions than answers. However, the faculty in the Divinity School could and should address and encourage other Yale faculty to consider two areas. First, how do we understand what it means to be a human being? This became a *cause célèbre* when Blake Lemoine, a Google engineer, hired an attorney to defend the rights of LaMDA, which claimed to be a person. The central—although not the only—issue is how we understand the human mind. Is it fully replicable in a machine? At present, LLMs are capable of processing vast amounts of data—far more than a human can and much more quickly—but it is not clear that the machines comprehend the text that they produce. Rather, LLMs predict what word is most likely to follow another based on an

algorithm that processes the available data. Christians, Jews, Muslims, Platonists, and others have for centuries held that human beings are capable of two different forms of thought: discursive rational thought (extending known propositions through the application of inferential rules) and intuitive thought (participating in a higher, nondiscursive form of knowing by which ultimate realities are immediately and reliably grasped). What would it mean for AI to become capable of these forms of thought? Or again if, in the future, AI surpassed human intelligence and was more self-controlled and disciplined in its processes, should machines become the decision makers for society and, if so, are there any limits to the spheres in which they make decisions? Or again, will AI change how we relate to ourselves and to one another? Will we offload our reading, thinking, writing, and analysis onto AI in ways that progressively erode our capacity to focus, think critically, and understand? Similarly, will we develop relationships with advanced machines rather than with human beings? Will we prefer simulated rather than human emotions? These are some of the anthropological questions that AI poses.

Second, how should ethical reflections shape and constrain the development and deployment of AI? When OpenAI began, it had an ethical board that met once, but not again. Our fear—shared by many—is that the market or international competition will be the sole drivers of AI development. We realize that AI research is being done internationally and that we cannot control what happens in other countries; however, this does not excuse us from setting some ethical controls in place. We need to think through the ethical implications and unintended consequences of AI research. For example, how do we account for bias? Recent experiences with Google Gemini have alerted us to the political biases that can be programmed into AI. Should AI be programmed idealistically—and who determines this? Or should it reflect the world as it is? Or again, how will we protect privacy, especially when consent to data use is a precondition for the employment of an AI program? Then, too, what values and aims should be programmed into future beings with artificial general intelligence? Should AI bots only "care" about human interests? Or should they care for the interests of nonhuman animals as well, even when the interests of these animals and the interests of human beings do not align? Should AIs care only about contemporary human beings or future generations as well and, if the latter, how far should this extend? Or again, what about the meaning of work and leisure? What happens if AI makes many occupations obsolete? Can humans find fulfillment in leisure in the same ways that they have found it in work? Can the traditions about rest within the Abrahamic faith traditions that speak of rest in idealistic terms offer any assistance? We could easily expand the list of questions, but these illustrate some of the concerns.

We are setting out to explore these theological and ethical concerns through the research and teaching of several faculty. Three members of our faculty (Jennifer Herdt, John Pittard, and Kathy Tanner) are in the initial stages of planning a conference on AI and theology. We hope that the proceedings will be published. The same three faculty are also planning courses: John Pittard has already begun including AI in his philosophy of religion courses and is planning a course that will address AI more fully from the perspective of the philosophy of mind; Jennifer Herdt and Kathy Tanner are planning to co-teach a course that will explore both ethical and theological issues. Other faculty are experimenting with the pedagogical values of AI. At least two faculty have asked their

classes to pose a stipulated question that touches on the core of the course that they are teaching to ChatGPT and then to write a critical evaluation of ChatGPT's response based on their knowledge from the course.

We believe that the Divinity School has something of immense importance to contribute to discussions about AI. One does not need to be personally religious to appreciate the impact of the school's potential contributions. In 2020, more than 350,000 communities of faith were in the United States. While the percentage of individuals who do not have affiliation with a religious tradition has grown notably in recent decades, religion remains a robust and potent force in the U.S. and in the world and will shape how many people react to and think about AI. Madeleine Albright was invited to give a lecture at the Divinity School after concluding her service as U.S. Secretary of State. She was asked what she had thought about religion while in office. She responded: "I didn't." However, after her lecture, she authored a book reflecting on the importance of religion and foreign policy entitled *The Mighty and the Almighty: Reflections on Power, God, and World Affairs* (2006). In a subsequent interview with CNN, Madame Albright explained why many thought religion was irrelevant to U.S. policies and offered a rejoinder: "Well, if we don't believe in the convergence of church and state, then perhaps we shouldn't worry about the role of religion. I think that we do that now 'at our own peril.' " Religion can be responsibly or irresponsibly practiced with significant consequences. We think that we should contribute to discussions about AI both at the level of the university and in training leaders of communities of faith who will shape how their community members engage AI.

David Geffen School of Drama/Yale Repertory Theatre AI Report

MARCH 18, 2024

Respectfully submitted by Shaminda Amarakoon, Joshua Benghiat, James Bundy, Anna Glover, and Tom Sellar

David Geffen School of Drama (DGSD) and Yale Repertory Theatre train and advance leaders in every discipline of the theatre, making art to inspire joy, empathy, and understanding in the world.

AI intersects with our work at many points: from technologies already embedded in or to be developed for creative processes, management, and pedagogy; to issues of intellectual property and diversity, equity, inclusion, and belonging (DEIB); to modes of performance and, indeed, as a subject of artistic expression.

Theatre makers—especially at a conservatory within a major research university—are likely challenged to compete for relevance in AI conversations for two reasons. First, by its nature, theatre is an accretive and traditional art form, slow to incorporate new technologies. It contains a large repertory and range of historical practices to comprehend, and theatre depends fundamentally on the embodied presence of human actors, crew, and audiences. Second, undercapitalization of the art form is likely to create a barbell distribution of innovation, in which AI technology for theatre will be advanced by and in the proprietary hands of large commercial entities, and artistic expression about or through AI will be largely the province of the most nimble creators working with a relatively small number of collaborators.

Still, each of these concentrations presents an opportunity for the Geffen School and Yale Rep to create partnerships with leading innovators by acquiring access to new technologies and/or programming artists who center AI as a subject of or technology for expression.

Key questions for the Geffen School are these *(with preliminary answers in italics):*

A. How can Yale lead in the understanding and protection of intellectual property rights of its faculty and students—and, indeed, of all artists, authors, and scholars with regard to AI? As the MacArthur fellow and Yale alumna Annie Dorsen has said, "let's be clear: these technologies were not designed to assist artists, they were designed to replace them."

   DGSD maintains an annual course, Law and the Arts, which draws the most students from other units of any course in the Geffen School and has begun conversations with Yale Law School about building collaboration among artists, managers, and legal scholars to address specific issues relative to AI.

B. How can Yale lead by addressing issues of intellectual and social responsibility, both for what goes into machine-learning systems and for what they produce: Who is responsible for deciding what is mis- or disinformation and what the systems are "taught"?

Who adjudicates what is biased thinking — or conspiracy theories? Who informs the machines which version of history is "true" and which is false — or how to mediate disputes over truth?

DGSD lacks broad expertise to address these questions and will seek further guidance from colleagues across the university, including those with expertise in computer science, ethics, history, and sociology. With respect to artistic programming, both the Geffen School and Yale Rep will take active interests in programming theatre works addressing such questions, as the field turns explicitly to the similarities and differences of human and artificial intelligence and the risks and rewards of the latter.

C. As AI technology develops, how can we deploy it to enhance safety and artistry? Lighting, sound, projection, and scenic automation systems may be linked in increasingly creative and efficient ways as AI learns to speak with all of them. Imagine a director who says: "I want to see one set of walls fly out, while another tracks on, while the turntable rotates, and all while the lights, projections, and sound are timed to the movement," and an AI system that accomplishes that programming in a matter of seconds.

   Therefore, can we invite companies to pilot the intersection of AI and our production technologies here at Yale? Is there potential to establish iterative or permanent theatre technology labs at the Geffen School with benefit for student artists, faculty, and the companies themselves?

   The new Dramatic Arts Building is programmed to contain six laboratory spaces: automation, mechanical, rigging, Mac, PC, and lighting/projection/sound. Design and technical disciplines are already engaged with technologies driven by AI, and we will continue to prioritize capital investment in software and hardware giving students access to state-of-the-art technologies. DGSD's innovation fund can and will be made available to support future residencies of and instruction by technology leaders.

D. What can we do to level the playing field among students with the resources to subscribe to ChatGPT, for instance, and students unable to do so?

   The Geffen School is evaluating providing subscription access to ChatGPT for all of its students, beginning in FY25.

School of Engineering & Applied Science and the Yale Taskforce on AI

APRIL 1, 2024

Preamble

While schools and units across Yale University grapple with how to harness and leverage AI in their work, the School of Engineering & Applied Science (SEAS) has a unique frame of reference for the modern sea change in artificial intelligence (AI). As the school that has been responsible for AI development to date, we have witnessed and responded to the tidal cycles of AI over its generations of evolution, and we occupy a critical position at the university to assess its technical foundations, veracity, and applicability in a principled manner. Engineering must leverage this perspective to serve the university as called for in our 2021 Strategic Vision, where we differentiated it from peers as the *school of engineering most integrally engaged in its larger university mission*.

This Strategic Vision likewise identified AI as the top priority area for the school, even at that time. To quote the 2021 document, since Yale has "an environment where cross-disciplinary collaboration is encouraged and impactful, Yale is uniquely positioned to take AI into its next phase. Doing so means tackling questions that lie precisely at the intersection between fields — often disparate ones, both from a technical and a cultural perspective." The advent of generative AI has only reinforced this need, with the increasing temptation to employ AI tools "off the shelf" rather than tailoring core AI techniques to exploit domain-driven potential.

What distinguishes this most recent rising tide? New technologies driving natural-language interfaces and "transformer models" have driven AI into a new consumer paradigm, and the unique capacity and architecture of graphical processing units (GPUs) has fueled an arms race in the space of computational capacity. Academia, not being in the business of meeting consumer demand, has been slow to pivot, forcing it to cobble together ad hoc solutions. At the same time, the power of these new technologies is accompanied by spectacular (and embarrassing) failures on simple tasks, euphemistically called "hallucinations," and a general unease with the reliability and factuality of the most popular systems deployed to date. The uncertainty latent in generative AI systems cannot currently be quantified, rendering these tools unreliable for direct application in science and engineering as well as other fields that require direct certainty.

As a result, engineering (and computer science in particular) finds itself in a critical position at the university. For Yale to lead in AI, a strategy that incorporates the broad strengths of the university with principled technical expertise and research leadership must be articulated and employed. Industry leaders recognize that institutions like Yale provide rich sources of domain expertise and high-quality data in addition to superlative talent, and we will benefit from an open posture toward industry efforts.

SEAS's AI initiative, articulated in its Strategic Vision, argues for emphasis on recruiting expertise in foundations and applications of AI. We hope this emphasis can serve and drive forward the goals

of the Yale Task Force on Artificial Intelligence. Cross-disciplinary recruiting already underway through multiple SEAS departments should be accelerated to this end. Such recruiting should enable and accompany appropriate industrial collaborations, efforts to obtain federal funding, and academic programming, also called for in the Strategic Vision.

Engineering research in AI and machine learning (ML) already interpenetrates a breadth of disciplines at the university, from economics, to law, to neuroscience, to arts and media. Activating and accelerating these partnerships will broadcast Yale's unique position to lead as a true liberal arts university, leveraging AI to advance the broad strengths of Yale, whether in research, recruiting, or educating our talented students for the technology-enabled society of the future. Yale Engineering sits ready to serve and help shape the university's strategy in AI.

Vision

While AI may draw its inspiration from nature, it is by definition engineered; hence, like any engineered system, its properties can in principle be engineered as well. Since activity in AI is wide ranging, we categorize AI into four major areas: (i) foundations (F), that advance the core discipline of AI; (ii) application-driven research (A), which co-evolves AI techniques using domain-specific knowledge; (iii) interactive AI (I), which focuses on the interactions between AI and society as well as the physical world; and (iv) users, where individuals use off-the-shelf AI tools to advance their work.

For Yale to lead in AI, robust activity will be required across foundations, application-driven research, and interactive AI. Our students require foundational knowledge in AI so they can be successful leaders. Since the foundations of AI are engineering-driven, we envision a hub-and-spoke model of activity in AI, with the "hub" centered in SEAS and the "spokes" spanning the breadth of the Yale campus. Engineering will provide the core training to our students by broadening its offerings to make them accessible to non-STEM disciplines (e.g., the new, Spring 2024 course CPSC 170: AI for Future Presidents). The Yale Center for Research Computing can host expert staff programmers, perhaps situated among applied mathematics or computer science, to assist the users of AI tools on campus. Engineering will serve as a natural interaction point with industry (e.g., the recently announced AI Alliance) to amplify impact.

Research Activities and Trends

The following offers a forward-looking perspective on research projects in engineering related to AI. We use **F** to indicate foundations, **A** for application-driven research, and **I** for interactive AI.

Large-language models and natural-language processing (F): Large-language models (LLMs) and natural-language processing (NLP) are among the most critical areas within AI. The Natural Language Processing Lab (Yale NLP), led by Arman Cohan, focuses on core science and technologies surrounding language models. This includes:

1. studying mechanisms of generalization in LLMs to understand to what extent they memorize knowledge versus their abilities to reason about and generalize to unseen scenarios;

2. developing improved algorithms for LLM training and inference to make them more efficient and reliable;

3. developing open source and transparent LLM technologies; and

4. efficient extension of LLM capabilities to novel scenarios and domains, including applications in science and engineering domains, to improve workflows.

Social robotics (A and I): Yale Engineering is well known for social and embodied artificial intelligence. Critically, Yale engineers will advance robot autonomy in order to enable new robotics applications that help people in positive ways. For example, Marynel Vázquez's group developed the Shutter robot as a flexible platform for studying fundamental problems in human–robot interaction. She uses this platform in her courses to teach students about perception and decision-making algorithms and to explain to the general public how AI algorithms work. Another example is the robot Ommie, designed to guide people through deep-breathing exercises and help them reduce anxiety. This project, led by Brian Scassellati, represents an important collaboration with the Yale School of Medicine.

Intelligent autonomy (A and I): The next frontier of AI systems will traverse the physical world to support humanity in solving real-world problems outside the lab. The biggest challenge facing the development of such AI in the physical world is the need for internet-scale data to effectively train the AI systems. Today, engineers are required to collect, curate, and siphon large amounts of data to produce what we currently see in AI—an unsustainable effort. Ian Abraham's research explores mathematical principles and algorithms for deploying reliable intelligent robotic and AI systems. His group investigates algorithmic reliability and formal guarantees of physically embodied AI systems through the development of optimal control and learning theory. These algorithms enable AI systems to explore, become curious, and play through interacting with their environment to collect data to learn skills of dexterity and agility.

AI and the brain (A): The origin of AI is, in fact, neuroscience, and the study of the brain inspired the development of algorithms. Moreover, in recent years AI has lent a lot to the study of the brain and helped us accelerate the pace of learning predictive models from large-scale neural data as well as understanding principles behind brain functioning. We want to leverage these advances and understand the brain better as well as to further advance the state of the art of AI. From an engineering perspective, we are well set up not just to understand and probe these neural processes, but to build algorithms that work with the brain in the loop. Engineering faculty members working on this activity represent SEAS's engagement with the Wu Tsai Institute (WTI).

Formal logic and languages for AI robustness and accountability (F and A): Typical topics that ensure the correctness of AI-based systems are robustness, fairness, correctness, accountability, and explainability. These properties can be reduced to mathematical formulas, and there are tools that can reason about them automatically. *Neurosymbolic reasoning,* an emerging field of computer science, uses formal logic and automated tools to prove the correctness of those properties. To use robustness as an example: once proven correct, formal verification guarantees the absence of all

possible adversarial input attacks. SEAS is collaborating with Yale Law School on a new application of neurosymbolic reasoning, combining LLMs and formal reasoning to enhance legal reasoning. We are developing a new generation of chatbots, so-called lawbots, to assist users with legal inquiries. The hallucinations that bedevil LLMs can be circumvented by applying rigorous methods that provide formal guarantees that lawbots supply accurate responses.

Computer systems (F): Progress in computer systems is one of the key foundations of modern AI. However, astronomical growth in AI-model complexity has resulted in unsustainable compute costs for AI. This leads to two natural questions: (1) How do we codesign new hardware and software systems for more energy-efficient and sustainable AI? and (2) Are there better ways to build Intelligent systems by drawing deeper inspiration from nature? Yale Engineering is working toward these goals in many active projects. For example: (a) we are creating new software systems that enable the use of AI at the edge; (b) we collaborate with the Yale School of Medicine's Department of Neurosurgery on AI-enabled brain-computer implants that can lead to new treatment options for epilepsy, Parkinson's, and autism; and (c) as part of the WTI, we develop software and chips inspired by neuroscience that aim to reduce the compute costs for AI by two orders of magnitude.

Applied and computational math (F): Glaring examples of failures in factuality—politely called hallucinations—force us to ask, Where does generative AI work, and why? The applied math group, working with many collaborations university-wide, is uniquely well positioned to tackle this question. Emphasizing theoretical analysis and the practical development of algorithms for science and engineering, we are working hand in hand to develop new algorithms and new mathematics for scientific impact. Before we ask, How can machine learning, AI, and new algorithms for scientific applications work in these varied and novel contexts? we must first ask, *Do* they work? Can we design better algorithms for better science? Are we doing good science for AI?

Historically, Yale has been a leader in computational simulations of complex scientific phenomena. Our strength in applied and computational mathematics can be a differentiator in the realm of simulation and generation of synthetic training data, which is increasingly necessary to improve model performance. Such simulation tasks often do not rely on GPUs and represent an important contribution to AI research that does not require massive computational resource investment.

Amplifying digital humanities (A): Advances in AI/ML will extend and accelerate past and ongoing work among computer science, the arts, and the humanities. New ML methods are needed to identify scholar-specified interesting features in various media-generic LLMs or image searches, requiring a joint effort between the humanities scholars, who understand the domain, and computer scientists, who understand digital representations. Concretely, new ML will assist experts in transcribing large numbers of Beinecke Library manuscripts that have been digitized but currently can only be searched using metadata. Current LLMs only provide small improvements to transcriptions created by existing systems and fail to produce satisfactory results, illustrating the need to develop AI algorithms with domain-specific knowledge.

Federal Funding Engagement

Engineering faculty are also part of two National Science Foundation–funded AI centers, where their work is focused on the foundations of AI. The Center for Data, Algorithms, and Market Design at Yale (CADMY), a collaboration between computer science, economics, and statistics and data science, has recently been involved in obtaining a Multidisciplinary University Research Initiative (MURI) award from the U.S. Department of Defense with colleagues from other institutions. Yale is well positioned to compete for center and institute awards from the NSF and other agencies, but faculty will need support regarding space and infrastructure for such efforts.

Tech Transfer and Industry Engagement

An appropriate AI ecosystem at Yale will include a robust and supportive environment for research ideas to be commercialized outside the university. The Roberts Innovation Fund has just selected its second round of projects, many of which are AI focused, across a range of applications, from medical, to legal, to hardware and semiconductor design. These efforts in turn have attracted the attention of major AI companies, which are providing cloud credits for these startup efforts, with Microsoft recently suggesting they will fund up to $150 thousand worth of cloud credits per award. Google and Amazon-AWS have already been supporting Roberts awardees, and OpenAI has offered ChatGPT credits as well. NVIDIA, IBM, and ASML are eager to work with Yale engineers to develop new research projects, and we see AI verticals as well as digital twin simulations and synthetic data as key directions for further development. We aspire to use central space on lower Hillhouse Avenue to convene industry collaborations that will support fundamental research as well as tech-transfer opportunities.

Education

Engineering's core responsibility will be to make AI accessible to everyone on campus while also offering breadth and depth to our computer science and other engineering majors. We actively imbue many parts of our curriculum, such as computer systems and formal methods, with AI tools and techniques. We have introduced CPSC 170 (an introduction to AI for nonmajors), and we have augmented our curriculum to offer foundational AI courses for our students, such as CPSC 370 (an introduction to AI) and CPSC 381 (Introduction to Machine Learning), a useful intro to many 400-level courses on AI. We are currently modifying 200-level courses to separate content that is relevant to any STEM student. This activity is in addition to 400-level AI courses (e.g., robotics, neural-network hardware, vision, medical-data analysis, etc.). In the area of LLMs, we offer advanced courses, such as AI Foundation Models (CPSC 488/588) and Natural Language Processing (CPSC 477/577), which play a crucial role in equipping our students with specialized knowledge.

Artificial Intelligence at the School of the Environment

MARCH 2024

Research

The Yale School of the Environment (YSE) is in the early stages of integrating artificial intelligence (AI) into its research and teaching. The most common uses of AI currently include creating and error checking computer code in languages such as R and Python, drafting and enhancing text, and exploring new research topics. YSE is using a variety of AI tools including Scite, GitHub Copilot, and Consensus, although the most widely used tool is ChatGPT 4.0. These AI tools are used to distill complex themes from open-ended prompts; conduct analytical tasks, such as image processing; conduct literature review; explore methods; and generate ideas for topic analysis. These uses showcase the potential of AI to boost productivity across coding, task management, data analysis, and writing by tailoring content to a variety of audiences and purposes.

While most uses fall into routine categories, several faculty members and their postdoctoral fellows and students are involved in advanced uses of AI. Below we describe examples of these uses.

One example uses AI combined with machine learning to identify candidate molecules for safe chemical–design research. This example highlights work by the Green Chemistry and Green Engineering Lab and involves evaluating the relationships between physicochemical properties of molecules and a variety of hazard outcomes. Developing green materials with desired properties requires a lot of trial and error. However, using AI models makes it possible to reverse engineer to find what chemicals are suitable for optimal material quality and environmental footprint. This approach leads to cost savings and minimizes waste. This project began with research to compare toxic chemicals to help understand the molecular descriptors that define toxic and nontoxic chemical categories, but the research has progressed to identifying molecular descriptors of aquatic toxicity.

Another lab is using AI to study how plant water use and photosynthesis are affected by the structure of leaves. Powerful 3-D imaging, known as microCT, can enable the visualization of leaf tissues in incredible detail thanks to AI-assisted analysis, which can look across thousands of images through the leaf cross section. This enables more detailed comparisons of different plant genotypes than previously was possible, which is important for research on drought tolerance for crops to enhance food security in a changing climate.

A postdoc working with a YSE faculty member is using AI to analyze large qualitative datasets, such as social media comments, interviews, and academic literature to understand the trajectory of public sentiment concerning diversity in environmental decision making. What's unique in this case is using AI to detect and mitigate biases within their own datasets and algorithms to improve the fairness and representativeness of their research outcomes.

A completely different kind of a use focuses on gathering information about research on AI using bibliometric analysis. This work includes using natural-language processing of journal article abstracts. One of the outcomes of this research has been to highlight our very limited understanding of the environmental impacts of developing and using AI. Both are energy intensive and represent fruitful areas for future research at YSE.

Nongenerative AI is widely adopted by the YSE faculty, particularly among those who rely on analysis of remotely sensed data. For years, machine-learning methods and other AI elements have been integral to these analyses.

Teaching

Incorporating artificial intelligence into teaching at YSE is at an even earlier stage of development than it is for research. Many faculty members have not begun to think about how AI will influence their approach to teaching and student work. All faculty members recognize that doing so is rapidly becoming a priority. Those who have considered how to incorporate AI into their courses have chosen to actively permit students to use AI tools in their coursework as long as they are transparent about how they do so. This is an area that we expect to change rapidly in the next several semesters.

In addition to including AI in existing courses, it is broadly recognized that training for faculty and students is an important near-term teaching priority.

Resource Needs

Site licenses for advanced AI tools, seed grants and fellowships for research and teaching, and tailored training programs have been most widely requested in the YSE community. Given the wide use of AI tools in daily research and coding, faculty, staff, and graduate students have all expressed the hope of gaining site-license support to access more secure and advanced AI tools (e.g., ChatGPT, Claude.ai, Google Bard, Github Copilot, Google CoLab, Scite, SciSpace, and Consensus) or gaining access through Yale Software Libraries.

Seed grants are expected to support more AI workshops and training sessions on creating and leveraging existing AI models responsibly. In addition, seed grants can empower faculty working on integrating large-language models (LLMs) and other transformer models into research, especially in remote sensing, green chemistry, and climate modeling. One faculty member has already expressed interest in hiring a postdoctoral or postgraduate associate to take charge of this area of research. In addition to seed grants, graduate student fellowships for encouraging research on employing AI methods are also important; this can complement existing AI graduate student fellowships that mainly focus on funding research related to theoretical AI methods.

Support on whether and how to use AI in externally focused training programs is also critical; for example, a tailored AI for the climate change and sustainability program can be included in one of YSE's new online certificate programs.

Plans

YSE is actively planning to ramp up research activity based on AI tools. We aim to develop schoolwide literacy in AI over the next year through initiatives with varying levels of formality and involvement. Faculty and graduate student researchers will do "share outs" of current AI uses, and workshops are being planned to inform a data-science hackathon in April 2024. YSE is also considering short courses offered by outside experts, a professional skill module on generative AI to make use of LLMs and other generative AI methods, and the development of toolkits on AI to empower community groups. We will craft an AI-use statement adapted from standard language on best practices, which will be passed out in syllabi guidance to new lecturers and will be available to existing ladder faculty. This statement could include, for example, language around the academic integrity policy. One instructor has proposed an AI and Sustainability course to investigate the potential impacts of AI on sustainability, both as a research tool and as a physical system that consumes resources. Finally, the school will examine the feasibility of hiring an AI specialist as a lecturer, research staff member, and/or IT professional to support AI research and teaching.

Report Authors

William Lauenroth, *Joseph F. Cullman 3rd '35 Professor* (chair)

Jennifer Marlon, *Senior Research Scientist, Lecturer, and Director of Data Science*

Owen Cortner, *Postdoctoral Associate*

Hannah Wang, *Doctoral Student*

Dani Heller Zero, *Director of Faculty Affairs & Data Analysis*

Yale FAS Report to the Yale Task Force on Artificial Intelligence: AI and Humanities Research and Education

APRIL 1, 2024

Introduction

This report is the response to the provost's call to understand the implications of generative artificial intelligence (AI) and to develop a strategic plan for AI for research and education in FAS Humanities.

Humanists study human beings in their engagement of the world—what human beings create, survive, engage, and discover. This scholarly work is also purposefully focused and slow. To move through the available evidence demands that people reflect with care, precision, and patience. Every humanistic tradition includes the exploration of new technologies and the social, spiritual, and material consequences of technological change.

Artificial intelligence presents a new and important chapter in the history of human technological development. The judgments and decisions that will shape our human future require consideration of social benefits and costs, of material advantage and disadvantage, and of security, wealth, and well-being.

As daunting as the technical questions are for fulfilling the vision of an AI-driven world, it appears that developing the ethics to govern innovation will be even more difficult. How should we manage technology—how do we shape outcomes, processes, and consequences—to ensure that human society is not only sustainable but also thriving? How will human beings create a future they want to inhabit rather than one defined by cataclysm and inhumanity? The answers to such questions are not simply technical; they are also profoundly humanistic.

Given that the FAS Humanities division comprises more than three hundred faculty and twenty-six departments, the summary is necessarily incomplete and subject to near-term changes as the rapidly developing technology is implemented. Some common themes arose during discussions among the faculty panel that drew from the many subspecialities in the humanities.

Key Findings

- The increasing prominence of AI in different areas of social life over the last few decades has been accompanied by explosive growth in the infrastructure required to produce, manage, and process such data.
- AI developers and investors advertise the individual consumer's efficiency gains, focusing on how personal tasks can be accomplished more efficiently with AI. But the broader social consequences for the emergence of AI infrastructure have been decidedly mixed.

- Data appropriation and processing are themselves a new type of resource extraction with global significance. Racial inequalities found in artificial intelligence practices are inherent in the data practices they analyze.
- The human labor necessary for data production requires its collection, curation, classification, labeling, and verification. Many forms of work ostensibly guided by AI require underpaid human workers performing rote tasks to shore up the impression that machines can do the work. AI promoters render invisible the labor costs of AI.
- Many scholars have explained that AI technology does not emerge from ethical industrial relationships. The development of AI tools regularly requires human rights violations, especially in the Global South. The long history of big data is by necessity one of invisible labor, in which the freedom or autonomy of participation of persons from whom data is generated is too often taken for granted.
- Big tech firms are incentivized to make AI something people buy. Any university should be wary of expending their resources on technology whose impacts are not yet adequately researched.
- Available research about the scientific uses of AI is cautious, even pessimistic about the kind of science that will emerge from AI-tool use. Novel tools and techniques are always prone to create monocultures. Knowledge-production systems that lack diversity in ways of knowing will be more vulnerable to errors and missed opportunities.
- Any corporate partnerships with Yale should have distinct guidelines in which university leaders prioritize the academic mission in any adaptation. Such guidelines have yet to be developed at Yale.
- There is a large body of best practices established about the engagement of new technology. Yale should be learning from the lessons derived from science and technology studies.
- While the potential benefits of AI are worth taking seriously, it is critical that scholars and developers of AI tools also consider the possibility that AI tools may limit rather than enhance understanding.

Opportunities

- The ability to collaborate and bring together expertise across the university offers a distinct advantage. Yale should find ways to encourage and reward those who can initiate such collaborations and should bestow advantage upon proposals in which humanists are included at the ground level to assure historical awareness and ethical consciousness.
- Librarians have begun grappling with the ethical nature of AI development and with the imperative of structuring a viable and sustainable future for delivering information. However Yale decides to invest in AI, it should simultaneously engage the libraries' significant expertise on data delivery.
- A cross-divisional certificate in critical computing has been proposed and would allow undergraduates to think about computer technology while considering ethical inquiries. Yale should support the expansion of certificates to welcome cross-disciplinary student advancement.

Yale FAS Report to the Yale Task Force on Artificial Intelligence: AI and Basic Sciences Research and Education

APRIL 1, 2024

Introduction

This report is the response to the provost's call to understand the implications of generative artificial intelligence (AI) and to develop a strategic plan for AI for research and education in FAS sciences. Given that the FAS Science division comprises more than 300 faculty and 8 departments, the summary is necessarily incomplete and subject to near-term changes as the rapidly developing technology is implemented in increasingly more powerful and novel ways of doing science. Nevertheless, some common themes arose during the discussions among the faculty panel that drew from the physical sciences, the life sciences, and their areas of overlap.

Key Findings

Common themes across the sciences:

- Data-driven and computational approaches to science have a long and rich history in the physical sciences, life sciences, and mathematics. This extends back to the 1800s for the use of applied statistics, but really blossomed with the use of computers to classify and automatically correlate data in the 1950s and later.
- The approaches to doing science using automated learning are very broad and rapidly are becoming more so. Computers have been excellent for organizing, correlating, and mining datasets. With modern AI tools, they can make original discoveries of new materials, find new behaviors of organisms or physical systems, resolve previously unsolvable problems like protein folding and protein evolution, produce new machine-checkable proofs that closely resemble human-readable proofs, discover new physical laws, and emulate aspects of human consciousness.
- The advancements alluded to above are coming more from machine learning (ML) than generative AI. The next phase of evolution in science will be with AI generating ML algorithms.

Opportunities for Increasing and Improving Research

- The ability to collaborate and bring together expertise across the university offers a distinct advantage. Yale should find ways to encourage and reward those who can initiate such collaborations.
- We must also find new ways to collaborate with industry, government, and academic partners at other universities. Industries and government agencies/laboratories do not always recognize the importance of using the domain expertise of university faculty in making progress that benefits the general good. Market forces and rampant competition can miss, or at least delay, significant advances. Yale should explore how to enable more collaboration with industry and government.

- The academic workforce is not sufficiently advanced in ML expertise. Moving from AI/ML accelerating research to AI/ML transforming research should be the university's goal.
- The shortage of graduate students, postdoctoral fellows, and research-rank faculty trained in AI/ML will only increase with time unless we move to provide training and opportunity for practice in AI/ML to postdoctoral fellows, graduate students, and undergraduates. **This training must be in the context of the scientific disciplines, not in replacement of or in simple addition to the training we already provide.**
- Additional training requires additional financial support: i.e., the academic workforce must expand to make it feasible to take on new capabilities. Faculty are too overworked as is, and few federal grants currently support this push into new frontiers while transitioning from present means of doing research.
- The current external funding environment results in a shortage of computational resources, particularly site licenses, purchased datasets, human support of computational tools, and sheer amount of accessible computational time.
- Increasingly, faculty have to be creative in leveraging other institutions to increase access to computational resources. Even the current subsidized cost of high-performance computing on campus is insufficient for an increasing set of scientific explorations.

Opportunities for Education

- Provide training for faculty in the use of AI tools.
- Encourage the creation — or the reimplementation — of courses that can apply and analyze AI in existing curricula.
- Encourage the adoption of curricular reform that responds to the increased availability of massive amounts of information from a wide variety of sources, some of which are more trusted than others.
- Encourage the practice of communication through human-generated text. Students need to be taught to generate *new* knowledge while simultaneously learning to think *independently.*

Concerns and Conclusions

Concern arises from possible steps in which computers do not amplify the capabilities of the human mind, but outstrip them completely by combining creativity with speed and the ability to use simultaneously massive datasets and previously gathered knowledge to reach results that are false. This is described as "hallucination" within the AI community, but it is a real concern for researchers or students who accept results because they "look" correct or authoritative. At the moment, none of the fields of science at Yale anticipate this to be a huge problem, but it is a worry that this may become an issue quite soon. It should be noted that this "false-information threat" is potentially far more damaging in the sciences than the prospect that malevolent actors will create something dangerous. Society depends on an uninterrupted stream of true results and maintained trust in the basic sciences to continue the applied research that undergirds our economic prosperity.

In addition, the following recommendation for university-wide initiatives and investments from the Yale School of Medicine (YSM)'s AI report includes ideas that FAS Science enthusiastically supports and would utilize.

- Increase computational power through acquisition of graphics processing unit clusters and cloud time as well as through bidirectional partnerships with other academic institutions, industry, governmental, and nongovernmental organizations.
- In collaboration with Yale New Haven Health System (YNHHS) and YSM, establish a robust HIPAA-compliant infrastructure to support the AI lifecycle, ensuring a secure computing environment *that is accessible to users university-wide*.
- Leverage the convening power of the provost's office to create spaces for cross-disciplinary collaboration in AI.
- Increase the number of computational graduate students. This should include increasing the number of students in Computational Biology and Bioinformatics (CBB) but should also include cross-training graduate students in the biological sciences in computation and students in the computational sciences in biological science. It is recommended that the Combined Program in the Biological and Biomedical Sciences (BBS) conduct a curriculum review and consider integrating teaching in AI across domain tracks.
- Consider opportunities for cross-courses or dual majors in computer sciences and common premedical majors. Yale College should consider AI courses for premedical students.
- Codify standards for scholarly contributions to multidisciplinary work in AI into appointment and promotions policies.
- Develop academic homes for those engaged in ethics related to AI.
- Enhance expertise in AI within Yale Ventures.

Yale FAS Report to the Yale Task Force on Artificial Intelligence: AI and Social Science Research and Education

APRIL 1, 2024

Introduction

This report summarizes ongoing discussions in FAS regarding the opportunities and challenges for social science research and teaching at Yale created by the emergence of artificial intelligence and machine-learning technologies. FAS Social Science is comprised of two hundred–some faculty in seven departments and six interdisciplinary programs and departments. Although there is no single conversation around, or approach to, AI among social scientists at Yale, the division as a whole is nonetheless deeply engaged with AI as a tool for research and teaching and an object of study. It is far too early to make a definitive assessment of its importance for the research and teaching mission of FAS social sciences, but it is abundantly clear that AI's potential and limitations are both easily exaggerated. That said, our conversations have yielded clear insights about areas in which AI is already reshaping research and will likely continue to do so in the coming years. Our report summarizes these findings and makes recommendations on what is needed for Yale to exploit and lead this innovation in the social scientific disciplines.

AI and Social Science in Three Dimensions

This section describes three dimensions for which AI and AI-related areas such as machine learning interact with the social sciences.

AI as a research and teaching tool

Emerging AI-based tools aid the research process across all disciplines including the social sciences. At a micro level, akin to a personal computer or the internet, these new tools can help with virtually every stage of the research process, including literature summaries, programming and coding, data collection, instrument testing, data analysis, presentation generation, and other routine—and not so routine—research tasks.[1] Researchers using these tools are likely to be far more productive than those who do not. For many social science faculty and graduate students, developing expertise in the use of AI tools, access to the best tools, and the computational resources to use them are essential to the quality and quantity of their future research. Our conversations indicated that the use of these tools is truly becoming widespread. It is not just statistics and data science (S&DS) faculty or psychologists affiliated with the Wu Tsai Institute (WTI). These may be particularly high-end and intensive users, but across the quantitative empirical social sciences, researchers are learning and using these methods (an Institution for Social and Policy Studies–led pilot program

1 See Anton Korinek, "Generative AI for Economic Research: Use Cases and Implications for Economists." *Journal of Economic Literature* 61:4 (December 2023): 1281–1317, for a review of how LLMs can be used in the research process in economics. Most of the assessments of what LLMs are good—and not so good—at doing for economics resonate with patterns we observe across the social science disciplines.

for a seven-week course for faculty about AI and large-language models (LLMs) had seventy respondents seeking twenty spaces).

The use of these tools in teaching is at an earlier stage. That said, faculty are experimenting with building interactive tutors tailored for their courses to provide new, more personalized resources to enhance student learning. Faculty are also incorporating the tools into their courses in the same ways they use other research methods in their teaching to allow students to learn by doing research.

AI as an object of study

AI and AI-related research areas such as machine learning are important areas of research in the social sciences. Faculty and students seek to understand the properties of AI methods and innovate in improving them. In some of this work, understanding AI itself is the end product of the scientific project. What are its properties? When and why does it work well? At what tasks does it perform poorly? But in some of this work, understanding AI is used as a means for understanding other complex systems. For example, work in cognitive science, linguistics, psychology, and S&DS, among others, seek to study AI to identify principles and form new hypotheses to better understand human cognition (and, in turn, use our enhanced understanding of human cognition to improve AI). Current research and future work in these disciplines is a two-way street of theoretical development and empirical discovery between neuroscience and AI. AI-related areas like machine learning also comprise a significant area of research for faculty in S&DS. Yale has recently added faculty who work in this area and it continues to search to strengthen this part of the department.

AI and society

Some technological innovations have dramatic consequences for individuals and societal institutions. They impact some individuals' jobs, the incomes they earn, the characteristics of the household unit, where they live, what they eat, who has political power, etc. AI is likely such a technology. Social scientists study individual human behavior and societal institutions and, as such, an emerging body of research across disciplines investigates the impact of AI on behavior and institutions. For example, Yale economists are developing theoretical and empirical frameworks for studying the impact of AI on labor-market outcomes across different groups in society. Political scientists are studying the effects of AI technologies on disinformation and political polarization. Although much of social science seeks to explain behavior and institutions, some of our social scientists study normative and ethical questions. Yale political scientists, sociologists, and cognitive scientists are developing ethical frameworks for thinking about how AI should work and how societies should regulate it.

Social Science Recommendations

This section describes several recommendations for investments needed to support FAS Social Science faculty and students (and social scientists across professional schools) in responding to the opportunities and challenges that AI is creating.

People

Our faculty conversations emphasized the importance of Yale investing in AI-specific human capital. This will involve both hiring new staff and faculty who have this expertise, but it must also include training our current faculty and students. While the need for AI experts who can train and serve as consultants for faculty and student research is likely one of interest across Yale's disciplines, the form of this expertise is likely to be somewhat domain specific. As such, we think targeting AI expertise for social scientists in FAS and across all the professional schools where social scientists work will strike a good balance between providing targeted relevant services and doing so at scale. This is, of course, exactly the sort of research infrastructure problem that the Data-Intensive Social Science Center (DISSC) was designed to solve, and our recommendation is that a significant investment be made in personnel for AI training (faculty and students) and consulting on specific research problems and projects. The demand for these resources is significant and we envision the need for workshops and consulting services launching as soon as possible, by fall 2024 we hope. At the same time, DISSC could foster intellectual community among faculty and students who are looking at questions about AI and its use and who can serve as a hub for cross-campus learning and exchange of ideas.

A different type of human-capital investment would entail postdoctoral fellows working at the very frontier of research on machine learning and AI. The Institute for Foundations of Data Science (FDS) has a new postdoctoral fellowship program that has successfully attracted top postdoc candidates. One possible vehicle for Yale to increase the density of frontier knowledge about AI and related areas would be to expand the program with a focus on identifying postdocs working in this area. FDS's mission is to advance research and to help scholars from other disciplines leverage that research in their work. It is well suited to help connect faculty across the university to relevant frontier research in this area represented by the envisioned new postdoctoral fellows.

While we do think that the importance of AI and AI-related areas suggests the need to hire new faculty working in these areas, our conversations do not propose the need for substantially different faculty hiring plans for the social sciences. The strategic plan for S&DS already includes a pillar focused on computation and machine learning.

The department will certainly be considering new developments in making these hires, but planned hiring provides adequate opportunities to bring researchers in this area to Yale. Other relevant departments and programs include cognitive science, psychology, and linguistics, all of which have either recently hired faculty working in these areas or have ongoing or planned future searches that will allow them to do so.

The final area of people to consider are graduate students. The Graduate School of Arts and Sciences (GSAS) has agreed in principle to allocate some fully funded graduate slots to combined PhD programs broadly in cognitive science. The cognitive science program, psychology, linguistics, philosophy, and computer science, which would have to be structured slightly differently, have not implemented this yet but should be encouraged to do so. The size and composition of funding

for S&DS's PhD program should also be reviewed with a focus on students working in machine learning and AI-related areas. The faculty-growth plan requires a similar increase in the size of the graduate program. (The complicated structure of graduate student funding in the department may also need to be reviewed in the context of any growth.) These areas of graduate student training have the greatest potential in the social sciences to contribute to AI research at Yale.

Computation and research infrastructure

To achieve Yale's potential to study AI and its effects in the social sciences, an investment in computational resources that is an order of magnitude beyond what we are doing currently is needed as are plans to sustainably support and adapt those investments. As the YTAI FAS Social Science/Jackson School of Global Affairs panel discussion highlighted, identifying what makes sense for universities to invest in and what should be left to industry in this space is a hard problem—one that we leave for the task force's larger conversation.

That said, we think neuroscience research in the social sciences is a good case study to help inform those discussions. At WTI, the Misha cluster is starting small and ramping up to try to best serve the diverse community of researchers doing computational work across the institute. Currently, Misha has 26 CPU nodes, 2 with large memory, and 16 GPU nodes with different grades of GPU. Storage servers totaling approximately 4 petabytes of storage are being added. They hope to grow the cluster significantly in the next couple of years, adding approximately 20 new H100 nodes. In order to keep up with advances in hardware, this equipment will need to be refreshed on an approximately five-year cycle. To give rough numbers, if the Misha cluster grows to 50 CPU nodes and 50 GPU nodes, funds will need to refresh on the order of $1.5 million per year (in 2024 dollars). An attractive model would be for the university to provide matching funds to researchers who raise funds from grants or gifts to purchase equipment.

Where does Yale stand relative to peer institutions with respect to computing resources for neuroscience? We can mention three data points. At Harvard the Kempner Institute was recently founded with a focus on artificial and natural intelligence that heavily overlaps WTI's mission. The Kempner's website describes an initial pilot installation of 144 A100 nodes, which was then expanded to include 384 H100 nodes. A recent senior faculty visitor from Harvard told us that, of the reportedly $500 million gift given to form the Kempner Institute, $150 million was allocated to computation. As a second data point, Princeton recently formed a new initiative called Princeton Language and Intelligence, which "seeks to develop fundamental understanding of large AI models; enable their application to research and education across academic disciplines; and study societal and ethical implications of AI as well as develop methods to avert any harms." Their website states that their GPU cluster will have 300 H100 nodes. A senior colleague at Princeton told us that the university has committed $20 million to build this cluster. Third, in January 2024 The University of Texas at Austin announced a new Center for Generative AI with a "Texas-sized GPU cluster" housing 600 H100 GPUs. According to these announcements, we lag behind our peers by a notable order of magnitude. This is much more relevant for maintaining our competitiveness for talent than

the roughly 3-log gulf with industry—Meta recently announced a contract with Nvidia to purchase over 600,000 H100 GPUs.

Finally, the expertise needed to train advanced computational models is as valuable as the hardware itself. This points to the need for a community of research software engineers to be shared across labs and centers, not only to amortize the cost of this expertise, but to create a community of non-ladder technical staff who can enable new styles of research within traditional academic environments.

Interdisciplinary Teaching on Artificial Intelligence and Emerging Technologies at the Yale Jackson School of Global Affairs

MARCH 2024

I. Executive Summary

The Jackson School's artificial intelligence (AI)–related activities are organized under the Schmidt Program on Artificial Intelligence, Emerging Technologies, and National Power. The program, begun in 2022, has established itself as a preeminent hub for scholars, practitioners, and students across Yale who are examining the array of complex global challenges associated with AI. The Schmidt Program's flagship yearlong course has attracted extraordinary student interest, and last year's graduates already have impressive accomplishments: founding an AI start-up company focused on commercial imagery analysis; publishing papers on AI-enabled disinformation; and launching student dialogues on the strategic implications of AI with Chinese and Taiwanese counterparts, to name just a few areas of notable achievement.

The Yale Jackson School of Global Affairs has appointed its first two inaugural Schmidt Senior Fellows for the 2023–2024 academic year, bringing both private- and public-sector AI expertise into the classroom. The Jackson School and the broader university now boast an expanded curriculum focused on AI and emerging technologies, led by Schmidt Program affiliates at Yale Law School and the School of Engineering & Applied Science. The Schmidt Program's AI Workshop and new Digital Ethics Workshop convene interdisciplinary scholars, who forge research collaboration across campus. Authors, policymakers, and industry leaders are regular participants in the Schmidt Program's Speaker Series, and these events now cater to hundreds of Yale students each week.

The Schmidt Program convenes scholars and practitioners to work across disciplines on the technological and strategic transformations that are reshaping our world. Since 2022, the multidisciplinary Schmidt Program has helped to launch seven new undergraduate and graduate courses dedicated to bridging the law, technology, policy, business, and ethics communities at Yale. It has brought prominent practitioners to campus as Schmidt Senior Fellows, offered postdoctoral fellowships to Schmidt Visiting Scholars, supported collaborative research, and developed a robust offering of lectures, workshops, and conferences to further the dialogue around emerging technologies and security studies.

Building on the solid foundation laid in the first year, the Schmidt Program now bears early fruits in research, teaching, and professional opportunities. It is poised to further enhance its impact at Yale and within its network of public, private, and nonprofit-sector partners in the years ahead.

II. Schmidt Program Personnel

Ted Wittenstein is the executive director of the Schmidt Program. He has built the program from scratch and teaches the core year-long course that underpins much of the program.

Senior Fellows this year include Beth Goldberg from Jigsaw and Andy Makridis, formerly from the CIA.

Beth Goldberg is the head of research and development at Jigsaw, a Google unit that explores threats to open societies. A Jackson alum, Beth now heads a team that investigates harms from disinformation and how to combat them.

Andy Makridis joined Jackson as a Senior Fellow after a 37-year career at the CIA. He was most recently its chief operating officer. Andy is a three-time recipient of the Presidential Rank Award, our nation's highest honor for civilian service.

III. Areas of Research Emphasis

The Schmidt Program has developed seven initial areas of research emphasis in close collaboration with affiliated ulty and practitioners. Each of these research areas is a focused teaching module within the Schmidt Program's new yearlong course and involves close collaboration with centers of excellence on campus and with non-Yale-partner institutions. In the years ahead, the Schmidt Program envisions supporting collaborative research, student internships, and other outstanding research proposals centered around these focus areas.

A. Cyberwarfare and the nature of conflict

Can there be deterrence and mutually assured destruction in cyberspace, or are these Cold War concepts not transferable to the current and future cyberthreat environment? How might military applications of AI augment cyberattacks or change the character of warfare? Are arms control and verification agreements possible with respect to cyberweapons or lethal autonomous weapons? Can we establish international laws and norms to discourage the proliferation and use of the most destructive AI-enhanced technologies?

Schmidt Program Partners

- Oona A. Hathaway, Gerard C. and Bernice Latrobe Smith Professor of international law; director, Yale Law School Center for Global Legal Challenges
- Andy Makridis, Senior Fellow, Yale Jackson School of Global Affairs; former chief operating officer, CIA

B. Disinformation and the future of democracy

How can AI tools help detect "deep fakes" and other forms of coordinated inauthentic behavior online? How can democratic policymakers and social media companies counter the threat of disinformation and online extremism without impinging on freedom of speech and other fundamental freedoms? What are the most effective techniques for educating citizens to the threat of disinformation, and developing counternarratives to combat conspiracy theories?

Schmidt Program Partners

- Beth Goldberg, Senior Fellow, Yale Jackson School of Global Affairs; head of research and development, Jigsaw
- Tauhid Zaman, associate professor of operations management, Yale School of Management

C. Competition and conflict in U.S.–China relations

How might AI alter the global balance of power? Are we headed for an AI- or tech-"Cold War" with China and, if so, what lessons can America learn from its generational contest with the former Soviet Union? What are the most effective regulatory approaches for monitoring imports and exports of sensitive AI technologies? How can policymakers decipher intent when certain AI tools are inherently dual-use? How might AI technologies proliferate to state and nonstate actors?

Schmidt Program Partners

- Arne Westad, Elihu Professor of History and Global Affairs; Director of International Security Studies, Yale Jackson School of Global Affairs
- Paul Tsai China Center, Yale Law School

D. AI ethics and safety

What techniques can help build AI systems that are reliable, transparent, safe, scalable, and aligned with human values? What ethical principles should govern military and intelligence applications of AI? How vulnerable are AI-enabled systems to subversion by malicious actors, such as through manipulating data inputs or "spoofing" images?

Schmidt Program Partners

- Luciano Floridi, founding director, Digital Ethics Center; professor of practice, Cognitive Science Program, Yale University
- Brian Scassellati, professor of computer science, cognitive science, and mechanical engineering; director of the Social Robotics Lab, Yale University
- Marynel Vázquez, assistant professor of computer science; principal investigator, Interactive Machines Group, Yale University

E. AI governance

How can social science research help public-, private-, and nonprofit-sector leaders make decisions regarding AI governance? What governance structures shape the development and deployment of machine-learning tools? Should there be a change in the current state of openness among the AI research community and, if so, how might that impact global prospects for cooperation or competition?

Schmidt Program Partners

- Information Society Project, Yale Law School

- Artur Pericles Lima Monteiro, Schmidt visiting scholar and lecturer, Yale Jackson School of Global Affairs

F. Nanotechnology and quantum computing

What are the implications for cybersecurity and global affairs of recent breakthroughs in both the theory and practice of quantum science? What are the potential roadmaps and notional timelines for the development of atomically precise manufacturing? Will these advances enable us to perform otherwise intractable computations, ensure privacy in communications, and develop new types of sensors and measurement devices? What are the most plausible malicious applications of these technologies by state or nonstate actors?

Schmidt Program Partner

- Steven M. Girvin, Eugene Higgins Professor of physics & applied physics, Yale University; member and founding director, Co-Design Center for Quantum Advantage, Brookhaven National Laboratory

G. Outer space

What vulnerabilities exist for space-based assets and related critical infrastructure? What form of legal system and regulatory regime is optimal to enable safe and secure long-term space development? What opportunities exist for adjusting existing treaties and norms to prevent an arms race in outer space as nations rush to develop antisatellite and hypersonic weapons? Can AI tools help detect and defend against these threats?

Schmidt Program Partners

- Kimball Smith Series, Department of Physics, Yale School of Engineering & Applied Science
- Yale Air Force ROTC Detachment 009
- Col. Todd Pennington, USAF, staff judge advocate, U.S. Space Command
- Gen. John W. "Jay" Raymond (ret.), former commander, United States Space Force

IV. New Academic Courses

Since January 2022, the Schmidt Program has developed or supported seven new courses for undergraduate, graduate, and professional school students focused on the vast array of law, technology, policy, and ethics challenges associated with AI. The Schmidt Program's signature new yearlong course, "Artificial Intelligence, Emerging Technologies, and National Power," is the connective tissue that strengthens collaboration with interdisciplinary faculty and practitioners across the campus and with institutions in the public, private, and nonprofit sectors. This cluster of courses now provides the foundation for how to develop technical fluency among aspiring policy leaders as well as expose STEM students to the legal, policy, and ethical dimensions of their research.

What follows is a summary of these classes. In the years ahead, the Schmidt Program aspires for these courses to form the basis of structured undergraduate and graduate academic programs, in which students may obtain a certificate or other form of distinction.

A. GLBL 6610, Parts I and II: Artificial Intelligence, Emerging Technologies, and National Power (Ted Wittenstein and guest faculty)

This yearlong graduate seminar—Parts I and II—examines how AI has the potential to alter the fundamental building blocks of world order. Machines capable of sophisticated information processing, toward the frontier of autonomy, pose tremendous opportunities for economic growth and societal well being. But the potential threats also are extraordinary: pervasive surveillance and digital authoritarianism; lethal autonomous weapons; AI-augmented cyberwarfare; sophisticated disinformation campaigns; and geopolitical instability as nations race to deploy these unpredictable technologies.

Drawing from a variety of interdisciplinary sources—and featuring guest scholars and practitioners from across the fields of computer science, data science, history, law, philosophy, physics, and political science—this course grapples with the challenge of building AI systems that are reliable, transparent, safe, scalable, and aligned with human values.

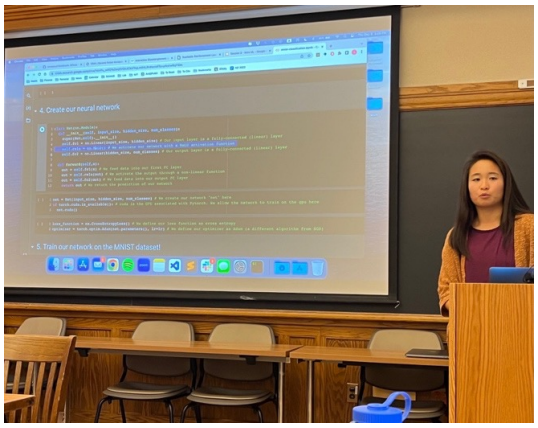
The seminar focuses on the Schmidt Program's seven core areas where AI and emerging technologies already pose significant security concerns: (1) lethal autonomous weapons and the nature of conflict; (2) disinformation and the future of democracy; (3) competition and conflict in U.S.–China relations; (4) AI ethics and safety; (5) AI governance; (6) nanotechnology and quantum computing; and (7) outer space development. For each of these subunits, the goal is to bridge the divide across the law, technology, policy, and ethics communities at Yale and to equip aspiring policy leaders with the requisite technical fluency to identify and respond to emerging threats and opportunities.

Over 170 students from various Yale schools and departments applied to the second iteration of the class. Students also traveled to Washington, D.C., to attend the Global Emerging Technology Summit on September 21, 2023, sponsored by the Schmidt Special Competitive Studies Project.

Part I of the course focuses on developing technical fluency and analytical skills. Students familiarize themselves with AI tools and consider the overall legal and policy landscape, learning from leading faculty and practitioners across Yale. As part of an innovative new curriculum design, for each of these sessions the featured guest presents and leads discussion for the first 45–60 minutes, while the second half of the 110 minutes is devoted to hands-on demonstrations, simulations, and group project work in the remaining time. Examples include:

- **Learning how bots work and utilizing AI to detect bots online,** *featuring Prof. Tauhid Zaman at the Yale School of Management and Pyrra Technologies.*
- **Building an algorithm for computer vision,** *featuring Prof. Brian Scassellati at the Computer Science Department of the Yale School of Engineering & Applied Science.*

- **Engaging in human–robot interaction,** *featuring Prof. Marynel Vázquez at the Computer Science Department of the Yale School of Engineering & Applied Science and the Yale Interactive Machines Group.*
- **Devising AI ethical principles and applying them to real-world risk-management scenarios,** *featuring former Yale World Fellow Sasha Brown at DeepMind.*
- **Learning and executing the quantum key distribution protocol,** *featuring Prof. Steven Girvin at the Department of Physics and the Yale Quantum Institute.*
- **Debating and revising Department of Defense Directive 3000.09, "Autonomy in Weapons Systems,"** *featuring DOD Principal Deputy General Counsel Corin Stone.*



Students write a reaction paper every third of the semester and organize into group project teams, where STEM and non-STEM students develop a research proposals that they then execute in the spring semester, applying what they have learned about AI to a concrete problem in global affairs.

Part II is devoted to further exploration of the international security implications of AI, which includes interaction with leading academics and practitioners along with a specific focus on U.S.–China technological competition. Students participate in the Yale Cyber Leadership Forum, a collaboration between the Schmidt Program and Yale Law School's Center for Global Legal Challenges, which brings together a mix of academics, lawyers, technologists, policymakers, and business leaders to discuss pressing cybersecurity challenges from different–vantage points. In spring 2023, the forum focused on the role of AI in Sino-American relations and whether the United States is headed for an AI- or tech-"Cold War" with China. Modules included partnerships with Yale Law School's Paul Tsai China Center and the Georgetown Center for Security and Emerging Technologies, and they focused on other strategic challenges, such as outer space (with U.S. Space Command, Blue Origin, and Voyager Space). A highlight of Part II was the Yale–Renmin Student Dialogue on AI, Emerging Technology, and U.S.–China Relations (discussed in detail at the end of this appendix).

Graduates of the first cohort have already made an impact. Founded by two alumni, a start-up company that leverages AI to improve disaster response sprang from class conversations and collaboration. Students have written rigorous scholarly papers. Others led student research

groups on AI and U.S.–China relations. Still others are now working in pioneering centers and government agencies.

The 2023–2024 cohort is poised for similar feats. Students in the current class already have founded a new start-up company that seeks to utilize AI to predict political violence while others have presented cutting-edge research on semiconductor geopolitics at leading academic conferences.

B. LAW 21023, Yale Cyber Leadership Forum—Bridging the Divide: National Security Implications of Artificial Intelligence (Oona Hathaway and Ted Wittenstein)

The Schmidt Program now sponsors the annual Yale Cyber Leadership Forum. In collaboration with Yale Law School's Center for Global Legal Challenges and the Yale Jackson School of Global Affairs, the forum brings together attorneys, technologists, entrepreneurs, policymakers, and academics to tackle the most pressing cyber challenges in a unique, interdisciplinary environment. The forum is codirected by Oona Hathaway and Ted Wittenstein and is team-taught by faculty across the university. The Yale Jackson School of Global Affairs and Yale Law School students can attend the forum in person for credit, engaging with speakers and writing reaction papers based on the sessions.

In spring 2023, thanks to the Schmidt Program, the forum focused on Bridging the Divide: Cybersecurity, Emerging Technologies, and U.S.–China Relations. It centered on a few questions and panelists offered a variety of perspectives in response:

- Is the United States headed for—or already entangled in—a tech-"Cold War" with China and, if so, what lessons can America learn from its generational contest with the former Soviet Union?
- What are the most significant challenges in U.S.–China relations today, and how do they manifest themselves in the cyber realm? How might trends in artificial intelligence and emerging technologies further amplify these tensions?
- What, if anything, is distinctive or unique about how China projects state power and influence online?
- What approaches to countering Chinese state-sponsored malicious cyber activity—across the full spectrum of diplomatic, law enforcement, economic, and military or intelligence options—have proven to be effective or ineffective, and why?

- What is the role of the U.S. private sector, and what mechanisms can better strengthen public–private-sector cybersecurity cooperation with respect to China?
- Despite geopolitical competition and differing conceptions of cybersecurity and Internet governance, can the United States and China bridge the divide to foster common cyber norms of behavior while avoiding miscalculation, escalation, and inadvertent conflict?

C. CPSC 611 / GLBL 6615, Topics in Computer Science and Global Affairs (Joan Feigenbaum and Ted Wittenstein)

The Schmidt Program has sponsored the first Yale course jointly offered by the computer science department of the School of Engineering & Applied Science and the Jackson School of Global Affairs. Now in its second year, this graduate-level seminar focuses on "socio-technical" problems in computing and international relations: challenges that cannot be solved through technological progress alone but rather require legal, political, or cultural progress as well. The class is designed to help bridge the divide across the law, technology, and policy communities at Yale, focusing on four key challenges at the intersection of computer science and global affairs: (1) disinformation; (2) cyberespionage; (3) encryption; and (4) artificial intelligence. The course is aimed at both STEM graduate students who desire greater exposure to the legal, policy, and ethical dimensions of their research and non-STEM graduate students seeking greater technical fluency. Students engage in interactive discussion, explore socio-technical challenges from diverse perspectives, and collaborate in interdisciplinary teams throughout the semester.

D. CPSC 310, Technology, Power, and Security: Political Challenges of the Computer Age (Joan Feigenbaum and Anat Lior)

In spring 2023, the Schmidt Program sponsored an undergraduate computer science seminar that examined the political challenges and opportunities associated with massive increases in the power of computational and communication technologies. Under what conditions do these technologies threaten citizens and governments, and how can leaders harness those technologies to solve problems? Case studies include cyberwarfare, cybercrime, the role of social media in democratic self-governance, authoritarian control, cryptocurrencies, and mass surveillance.

E. AI Policy: A Comparative View (Anat Lior)

In spring 2023, Schmidt Visiting Scholar and Lecturer Anat Lior offered a Jackson graduate-level seminar that explored how AI presents new and complex legal issues that current law systems are not well equipped to handle. Chief among those challenges is the "black-box" issue, or the opaque and uncertain way AI-based algorithms make decisions. This issue raises numerous legal problems, including accountability, liability, predictability, and establishing legal causal links.

This course delved into the way AI is constantly shaping and reshaping different legal systems around the world, concentrating on the U.S., EU, China, and Canada. It focused on AI's usage and influence upon different legal arenas, such as the criminal justice system, national security, and elections. To illustrate the legal shift AI has created throughout the world, the class reviewed three case studies focusing on facial recognition, the Covid-19 pandemic, and the Chinese social-credit system. The course examined different solutions suggested by scholars and government officials for AI governance to allow the wide usage of this technology while mitigating its negative effects.

F. The Global Law and Policy of Artificial Intelligence (Artur Pericles Lima Monteiro)

In spring 2024, Dr. Monteiro, Schmidt Visiting Scholar and Lecturer in Global Affairs, taught a course that explored how law made AI, what challenges it poses to leading products, and which governance options are now being discussed. Throughout the semester, students considered the geopolitics of the law and policy underpinning AI, with the dominance of U.S. law now facing a challenger in the "Brussels effect" of EU regulation. Students will discuss topics such as facial recognition, mis- and disinformation, workplace surveillance, privacy, and copyright.

Taught both for those interested in developing artificial intelligence systems and those who want to study its many implications, understanding the legal background against which AI stands is crucial. Legislation, policymaking, and case law have enabled the dominant business models for AI products, and they can also change as new winds prompt regulators into a reexamination of choices and incentives. Indeed, the law is not yet settled for central AI questions. Ongoing litigation, enforcement action, and debates threaten the lawfulness of AI models that implicate privacy and data-protection laws, copyright law, and antidiscrimination laws, to name a few. At the same time, new governance structures and regulatory frameworks are being proposed or taking hold. Students will develop a foundational grasp of legal concepts and policy issues that will benefit both those who see themselves as AI practitioners and those who will research or critique it.

G. AI for Future Presidents (Brian Scassellati)

During spring 2024, Professor Scassellati launched a new undergraduate lecture course, which introduced AI to non-STEM majors. This course is the outgrowth of Professor Scassellati's participation in the Schmidt Program's yearlong course, in which he has developed a four-part teaching module on AI fundamentals: (1) data and algorithms; (2) natural-language processing; (3) computer vision; and (4) robotics. This Schmidt Program module, in which students learn to

design their own algorithm for computer vision, is the foundation for scalable curricula that will reach hundreds of students in Yale College.

H. Generative AI and Disinformation (Beth Goldberg)

Inaugural Schmidt Senior Fellow Beth Goldberg, head of research and development at Jigsaw, has developed a course on generative AI and disinformation that wais offered in spring 2024. The course leveraged Goldberg's experience leading an interdisciplinary team of researchers who investigate online harms, from disinformation to violent extremism as well as her work at the U.S. Department of State, where she focused on digital security, investigative journalism and information access in authoritarian regimes.

V. Programmatic Activities

During 2023 and 2024, the Schmidt Program was able to host a range of high-profile conferences, workshops, speakers, and other events that enhanced the overall research and teaching agenda. This helped establish the program as a thriving intellectual hub for AI at Yale.

A. AI Speaker Series

The AI Speaker Series brings leading AI technologists and thought leaders to campus, strengthening Jackson's collaboration with centers of excellence across the university. Recordings of these sessions also continue to provide valuable teaching resources.

B. Global Emerging Technology Summit

In September 2023 and September 2024, the Schmidt Program sponsored 20 students in its yearlong course to travel to Washington, D.C., to attend the Global Emerging Technology Summit sponsored by the Special Competitive Studies Project. The full-day conference featured sessions with Senate Majority Leader Chuck Schumer, Dr. Eric Schmidt, Dr. Condoleezza Rice, and an impressive array of leading strategists and technologists.

C. AI Workshop at Yale University

As part of the Schmidt Program, this workshop is aimed at connecting anyone from the Yale community who is interested in the evolving topic of AI from any perspective. It invites students and scholars from the Yale community to participate in a new series of workshops on artificial intelligence and aims to provide in-depth engagement with works in progress.

The goal of the workshop is to establish an AI scholarship hub at Yale. It explores the regional, national, and international policy implications of AI. The workshop facilitates conversations on governance structures, institutional design, international relations, trade, and regulation of AI. The workshop encourages participation from all disciplines on legal, economic, social scientific, technical, regulatory, and policy issues related to AI in a global context.

Some of the subjects the workshop featured in its first year included AI systems applications in healthcare, geopolitics of AI governance, history of machine learning, AI and music, algorithmic fairness, and content moderation.

For the workshop's second year, more than 60 people across Yale joined the first session, including students and faculty from the Jackson School, Law School, computer science department, political science, the Yale Graduate School of Arts and Sciences, the Medical School, the Digital Ethics Center, and the Paul Tsai China Center.

D. Digital Ethics Workshop

This one-semester workshop was a collaboration between the Digital Ethics Center and the Schmidt Program. Professor Luciano Floridi, founding director of the Digital Ethics Center and professor of practice in cognitive science, introduced key topics in the field of digital ethics, such as the nature and ethics of AI, digital sovereignty, and the relationship between environmental and digital issues, among others.

The goal was to understand the epochal transformations affecting human self-understanding, the shaping of information societies, the conceptualization of reality, and the new forms of interactions among individuals and between humanity and the world, from an ethical, normative, social, and political perspective. This workshop enabled participants to develop a deeper and more critical understanding of the digital revolution and its ethical impact and implications.

E. Yale-Renmin Student Dialogue on AI, Emerging Technology, and U.S.–China Relations

This Yale-Renmin Student Dialogue explores how trends in AI and emerging technologies impact Sino-American relations. How can the United States and China build AI systems that are reliable, transparent, safe, scalable, and aligned with human values? What ethical principles should govern the deployment of these tools? How can the United States and China compete to develop AI and other emerging technologies while also avoiding miscalculation, escalation, and inadvertent conflict? In spring 2023, Yale and Renmin University faculty members moderated the dialogue and facilitated this virtual exchange, sharing their own perspectives.

Following the success of the dialogue, Ted Wittenstein was invited to visit Renmin counterparts in Beijing in summer 2023 as a prelude to the first in-person session of the dialogue, which took place in March 2024. The dialogue took place on the campus of Renmin University. Nineteen students from the Jackson School, the Yale School of Management, the Yale School of Engineering & Applied Science, and Yale College participated.

The Schmidt Program also has forged ties with Taiwanese leaders across the public, private, and nonprofit sectors, given Taiwan's critical role in the global semiconductor supply chain. Following their dialogue with Renmin University counterparts in Beijing, Schmidt Program students and accompanying faculty and fellows traveled to Taipei in March 2024 for a series of high-level discussions focused on AI, emerging technology, and cross-Strait relations. Members of the Jackson School delegation held discussions with Foreign Minister Joseph Wu, Minister of Digital Affairs Audrey Tang, experts from the Taiwan Center for Security Studies and Taiwan AI Labs, and a number of former government officials and party leaders. In addition, the Yale group visited the city of Hsinchu, a major hub for science and technology development. They met with leaders from Taiwanese tech companies that included Taiwan Semiconductor Manufacturing Company (TSMC), UMC, and MediaTek, as well as the Semiconductor Institute of National Yang Ming Chiao Tung University.

Yale Law School

Report for the Yale Task Force on AI

MARCH 18, 2024

Yale Law School (YLS) has been immersed in learning more about the power of artificial intelligence (AI) for more than a decade. Through intellectual centers and clinics, courses, workshops, and special events, the Law School is investigating and investing in AI for legal education and research, even as our scholars propose guardrails and regulation around its use. In keeping with our tradition of thinking bigger and bolder about emerging trends in legal education, we are teaching our students in a completely distinctive fashion. As Professor Scott Shapiro says, "At Yale Law School, we don't just teach students the law, we teach students how to teach AI models the law."

In recent years, we have added course offerings, reading groups, and seminars to address AI and other technology frontiers as issues of pressing legal concern. Courses include Liability and Regulation at the Frontier of AI Development; Artificial Intelligence, the Legal Profession, and Procedure; Cybersecurity; The Global Law and Policy of Artificial Intelligence; and Technology in the Practice of Law. The Lillian Goldman Law Library technology team offers workshops on Practical AI for faculty, staff, and students as well as individualized consultations on AI applications in empirical research.

Yale Law School's Information Society Project (ISP), a leading intellectual center in this space, has spearheaded many events focused on AI, including two AI governance symposia, and brought in experts from across the public and private sectors to educate and consult with students. YLS will host the Propaganda and Emerging Technologies Conference in April 2024; it will also host a conference entitled "The Normative Philosophy of Computing" next fall. The Tsai Leadership Program is poised to be a key funder of future workshops, reading groups, speaking events, and courses taught by top-flight experts on AI from around the world.

Professor Jack Balkin, a member of the Yale Task Force on AI and head of Yale ISP, illustrates what the future vision for legal education looks like. "If all goes well, we'll have many faculty members who incorporate AI issues into their scholarship. We won't have scholars who just do AI—we'll have scholars in business regulation, contract law, tort law, civil procedure, freedom of speech, copyright, etc., who have incorporated issues regarding AI regulation into their courses," he says.

As we build out these programs and resources, Yale Law School aims to use our leadership status in the world of legal education to drive the conversation forward to shape the future of AI. We will equip our students with the tools they need to lead and serve in a new world where AI is a significant component of their future careers.

"In the future, people will use AI regularly in business, legal education, [and] government," says Balkin. "People will find new uses for it, and it will shape the kind of skills we'll teach our students. AI will be integrated deeply into elements of the curriculum. It will also affect clinical education. AI programs will do a lot of things that are time consuming and repetitive, and students will regularly employ them. I expect that a decade from now, there will be significant integration of AI into the scholarly agendas of our faculty and in the everyday life of the law school."

"Conversely, what we do at Yale Law School will affect AI as well," adds Balkin. "YLS scholars will likely be on the front lines of developing legal solutions for the regulation of AI and related technologies, as well as adapting older doctrines and legal structures to account for AI. They will build new AI-based tools for legal research, using AI to ask new kinds of questions. Finally, they will create new methods for employing AI in legal education."

Through Yale Law School's AI programming, faculty and students are already pioneering how AI can be used in the law. "One of the things people always say with AI is that data is king and it's hard to get good data. Our students produce incredibly high-quality data that gets thrown away or forgotten," says Shapiro. But he added that data could become an "invaluable asset for the future of legal education and knowledge" if used to train models. "Our research is designed to see how we can leverage this new and extremely exciting technology to help fill gaps in access to legal services and prepare students for the transformation of legal practice," Shapiro concluded.

Current Work at YLS

Courses and clinics

YLS students are investigating constitutional and regulatory elements of AI in core courses. They are also learning how to hack and test AI and other emerging technologies for themselves, and to program AI models for use in clinics. The following is a list of courses, clinics, and reading groups at YLS that address AI.

- Law and Large Language Models (Scott Shapiro and Ruzica Piskac)
- International Technology Capital Markets (Keerthika Subramanian)
- The Information Society (Jack Balkin)
- Liability and Regulation at the Frontier of AI Development (Ketan Ramakrishnan)
- Torts and Regulation (Ramakrishnan)
- Technology in the Practice of Law (Femi Cadmus)
- Artificial Intelligence, the Legal Profession, and Procedure: Seminar (William Eskridge, Jeff Chivers, and Theodore Rostow)
- Law, Security, and Logic (cross-disciplinary course with computer scientist Ruzica Piskac)
- Cybersecurity (Shapiro and Sean O'Brien)
- Cybersecurity, Cyberwar, and International Relations (Ted Wittenstein)

- Global Law and Policy of Artificial Intelligence (Artur Pericles Lima Monteiro)
- Issues in Financial Regulation: Focus on Financial Technology (Saule Omarova)
- Media and Technology Industries: Public Policy and Business Strategy (Jonathan Knee)
- Social Media Governance (Tom Tyler)
- Private Law Clinic (Daniel Markovits and Andrew Miller)
- Media Freedom and Information Access Clinic (David Schulz and Jack Balkin)
- Advanced Media Freedom and Information Access Clinic (Schulz and Balkin)

Library

The Lillian Goldman Law Library is taking a leadership role in the use of AI for legal research. Law librarians incorporate AI into research classes, offer workshops and events aimed at demystifying AI, and investigate possible applications of emerging technologies in law. Femi Cadmus, law librarian and professor of law, teaches Technology in the Practice of Law, which utilizes AI tools.

- The Library Technology Team provides individual consultations for faculty and students, assistance with AI-related applications, and assistance with empirical and data projects
- Practical AI workshops for staff and students
- Courses that utilize AI tools include Advanced Legal Research: Methods and Sources; Research Methods in Statutory and Regulatory Law; Research Methods in Judicial History; and Research Methods in Foreign and International Law

Events/Conferences

Many years before ChatGPT became a household name, YLS had already taken a leadership role in investigating its potential. We regularly host top-flight experts for speaking engagements and discussions to educate our community on issues related to AI and the law. Major events hosted by the Tsai Leadership Program are planned for the coming year.

Tsai Leadership Program

The program's funding will enable visiting lectures and teaching opportunities from leading AI experts from around the world. A full slate of events for 2024–2025 is being finalized. One potential speaking event—still in the early stages of planning—may include the head of strategic partnerships on the Global Affairs Team at OpenAI.

- Regulating AI: A Conversation with Gary Gensler, chair, U.S. Securities and Exchange Commission, February 13, 2024 (cosponsored by Yale Law & Business Society)
- Planned workshops, guest lectures, and visiting faculty will tackle this topic and enhance the current curriculum

Information Society Project

Under Professor Balkin's leadership, ISP has been at the forefront of research related to technology and the law for decades. "To study platforms now is to study AI," Balkin says. Postdoctoral students with the ISP are already studying AI as part of their scholarship, and with funding from the Tsai Leadership Program, a dedicated AI research cohort will start at YLS in 2025. ISP has hosted numerous events. Here is a sampling:

- Unlocking the Black Box (April 1–2, 2016)
- We Robot (March 31–April 1, 2017)
- (Im)Perfect Enforcement Conference (April 6–7, 2019)
- Big Tech & Antitrust Conference (October 3–4, 2020)
- News and Information Disorder in the 2020 U.S. Presidential Election (December 4, 2020)
- Workshop on Private Law and Emerging Technology (September 3, 10, 17, 2021)
- AI Governance Virtual Symposium (series) (March 10, 2021–June 29, 2023)
- Technologies of Deception Conference (March 25–26, 2022)
- Data (Re)Makes the World Conference (March 3–April 1, 2023)
- AI Governance Virtual Symposium: Exploring AI Accountability Policy with Russ Hanser (June 29, 2023)
- Encoding Poverty: The Algorithms Too Few People Talk About, Amos Toh, Human Rights Watch (January 30, 2024)
- Developing AI Accountability Policy: A View From the Field, Ellen Goodman, Rutgers University (February 6, 2024)
- The Prediction Society: AI and the Problems of Forecasting the Future, Hideyuki Matsumi and Daniel J. Solove and *Yale Journal of Law and Technology* (February 20, 2024)
- Propaganda and Emerging Technologies Conference (April 5–6, 2024)
- Normative Philosophy of Computing Conference, Fall 2024 [UPCOMING]

Solomon Center for Health Law and Policy

- The Law and Policy of AI, Robotics, and Telemedicine in Health Care, Solomon (November 2, 2018)
- Legal, Ethical, and Equity Issues Surrounding AI in Healthcare (panel), Solomon and ISP (April 3, 2023)
- Generative AI and Medical Advice, Solomon, Claudia E. Haupt (November 8, 2023)
- The Transformative Impact of Artificial Intelligence on Healthcare and Legal Practice: A Craig Wasserman 86/Wachtell, Lipton, Rosen & Katz Alumni Breakfast in New York City (November 16, 2023)

Paul Tsai China Center

- The Frontiers of Artificial Intelligence Governance in the U.S. and China, Paul Tsai China Center (September 12, 2023)
- Responsible AI in Global Business Conference, School of Management, with in-kind support from Paul Tsai China Center and other partners (March 1, 2024)

Center for Global Legal Challenges

- Yale Cyber Leadership Forum, with the Yale Jackson Institute for Global Affairs, Oona Hathaway and Ted Wittenstein (February 4–April 29, 2021)

Faculty work

YLS faculty are at the forefront of investigating the possibilities of AI for use in legal practice as well as proposing guardrails for its use.

- **Jack M. Balkin**, Knight Professor of constitutional law and the First Amendment, is a leading expert on technology and constitutional law, and has been writing about governance of robotics and AI since 2015. The founder of YLS's Information Society Project, he also directs the Abrams Institute for Freedom of Expression, and the Knight Law and Media Program at YLS. Under Professor Balkin's leadership, in 2021 the Media Freedom and Information Access (MFIA) Clinic began hosting the Tech Accountability & Competition Project, which is supervised by David Dinielli. Funded by the Leadership Program, Professor Balkin intends to create an ISP postdoctoral program for J.D. graduates who want to teach law, specializing in AI, technology, and civil liberties.
- **Scott J. Shapiro**, Charles F. Southmayd Professor of law and professor of philosophy, investigates the intersections of law, philosophy and technology. He has written widely on cybersecurity and published *Fancy Bear Goes Phishing* in 2023. Through the Leadership Program, Professor Shapiro is planning an AI lab in which students, programmers, and computer scientists will train responsible AI models for use in legal contexts.
- **Robert C. Post**, Sterling Professor of law and former dean of YLS, specializes in constitutional law and the First Amendment. He is a trustee on Facebook's oversight board and is interested in governance of emerging technologies and implications of AI for intellectual property law.
- **Abbe R. Gluck**, Alfred M. Rankin Professor of law and faculty director of the Solomon Center for Health Law and Policy at Yale Law School and professor of internal medicine (general medicine) at Yale Medical School, leads the Solomon Center's groundbreaking work on technology and policy in healthcare.
- **Ketan Ramakrishnan**, Associate professor of law, teaches Liability and Regulation at the Frontier of AI Development and Torts and Regulation.
- **William Eskridge Jr.**, Alexander M. Bickel Professor of public law, teaches the course Artificial Intelligence, the Legal Profession, and Procedure and was the keynote speaker at the Artificial Intelligence & Chatbot Summit at the Washington, D.C., Bar in September 2023.

Yale School of Management
Report on Summary Observations and Recommendations for AI

APRIL 1, 2024

At the request of the provost, the Yale School of Management (SOM) provides the following preliminary overview of plans for AI as well as a summary of AI activities.

Overview of Plans for AI at SOM

As a professional school producing graduates for leadership in a wide range of industries, the Yale School of Management naturally has an orientation toward interdisciplinary thinking and application and must adapt quickly to changes in practice in order to maintain relevance. Current SOM AI activities span both research and teaching and touch all faculty groups at the school. (SOM faculty are organized into six distinct groups: accounting, finance, economics, marketing, operations, and organizational behavior with a diverse set of reference academic disciplines, mainly in the social sciences.)

A key observation about our work with artificial intelligence (AI), machine learning (ML), and related technologies is that they connect with a wide spectrum of topics related to business and society, which we categorize as follows:

1. The nuts and bolts of AI/ML, such as big-data analysis, software engineering/coding, database systems, etc.
2. Specific applications for AI/ML within organizations and fields of study (e.g., asset pricing, predictive decision making, social media analytics, etc.).
3. Macro-level phenomena related to AI's impact on markets, competition, regulation, the wor place, global workforce, and society.

The above categories of inquiry apply to both the research our faculty are engaged with and the courses they teach. The applied orientation of our research tends to create synergies with teaching that makes it less necessary for SOM to centrally direct the creation of new courses or course content related to AI. SOM faculty naturally develop lessons that draw upon their research insights, and SOM students require skills related to these innovations in order to prepare for careers in a wide range of functions and roles. We also make use of adjunct faculty to enrich our offerings in specific niches that practitioners, given their proximity to such niche areas, are best- suited to teach.

At the same time, we believe that, when it comes to research, market forces are strong. Faculty, particularly those in their early careers, and doctoral students have strong incentives to be on the leading edge of technological innovation in their respective disciplines.

Given these characteristics of SOM and its ecosystem, which includes competition from peer institutions and industry recruiters looking to hire from among our schools, we do not see a strong need to catalyze adoption of technology that is already occurring organically. This influences the degree of centralized planning we are engaged in on this front. Therefore, the two threads driving our approach to AI at SOM are (a) facilitation and support of the organic activities of our faculty in the areas listed above, and (b) facilitation of exchange of ideas and experiences across faculty.

Researchers in this space need access to computing resources, software tools, and research assistants familiar with the relevant technologies. We see our efforts as a continuation of investments we have made to support research using big data. See below.

On the teaching front, the immediate concerns center on policies and guidance regarding student use of LLMs in their work, academic honesty, and support for faculty who want to integrate AI tools into their pedagogy in innovative ways.

Summary of AI Activities

A recent survey of SOM's faculty provided us with specific information and feedback about how they presently use AI/ML in their research and teaching and what priorities we might identify for the school and university with respect to infrastructure and resource allocation. We refer to some of the survey findings in describing our current activities. (Our survey instrument appears in Appendix 2 and results are based upon 46 responses out of ~100 faculty surveyed.) Appendix 3 lists a sampling of additional AI-related activities that have recently taken place at SOM.

Research Activities

For the last few years SOM has been building out its infrastructure to support research that has become increasingly data intensive. This effort has included increasing the capabilities of SOM's internal research-computing cluster as well as expanded funding for research assistants (primarily predocs). These investments will naturally support research in AI. We also note that SOM's vibrant research seminar series—and, in particular, the school-wide internal seminar series—play important roles in exposing faculty to new techniques and research questions and lead to sharing knowledge across disciplines. This is essential.

Penetration of AI/ML is high in research at SOM. More than 50 percent of the respondents are using it "Somewhat," "To a Large Degree," or as "Core to what they do" in research. About 50 percent of the researchers also see themselves as developing new methodologies.

About 50 percent of those using AI/ML in their research are making use of cloud resources. The rest make use of the SOM research-computing cluster and/or the central Yale HPC. Faculty satisfaction with the resources available is discussed below.

Teaching Activities

Our survey shows that penetration of AI/ML is also high in teaching. More than half of the respondents indicated the amount they teach about these subjects as "Somewhat." In Appendix 1 we give a sampling of some of the courses we already offer in this space.

Supporting pedagogy at SOM is done in many ways, including faculty-development workshops, where we often partner with the Poorvu Center for Teaching and Learning. In April 2024, SOM will host its first formal faculty discussion on AI in coordination with Yale's Poorvu Center—to ensure that we are coordinating conversations and setting priorities that align directly with our faculty's expertise and interests.

Gaps and Looking Forward

Our survey found that 24 percent agree or strongly agree that Yale infrastructure is sufficient for their work, 62 percent neither agreed nor disagreed, and 14 percent disagreed or strongly disagreed that infrastructure was sufficient.

Delving deeper into the responses of those who expressed dissatisfaction, we found a variety of concerns that we believe are instructive in understanding faculty needs.

1. There are challenges using high-performance computing (HPC) in terms of complexity, support, and suitability. Cloud services have many advantages.
2. Using cloud service is expensive, particularly when training models. Some faculty are burning through research funds.
3. Some need support in terms of expertise that would come from research assistants or Yale staff, who can help move up learning curves and help with implementation. This applies both to teaching and research fronts; that is, some faculty know how they could use AI, and they want to, but they cannot do it by themselves.
4. Expertise or infrastructure is not yet in place at Yale to support faculty who are building innovative teaching tools based on AI. Thus, leading-edge users are hitting roadblocks.

The above comments raise important questions about the university's approach to make versus buy. Part of the equation also depends on how quickly adoption of AI will take place. The cost of this technology will decrease over time, and ease of use will improve, as it always does.

If faculty inject AI into their activities (both teaching and research) at a high rate, however, it will turn an increasing proportion of their activities into something computationally intensive. In the short run, we should expect this to outpace the benefits of Moore's Law and increase Yale's computing costs.

At this early stage it is difficult to state a long-term vision for these emerging technologies at SOM. But we are as always committed to broadly supporting faculty-driven innovation and continue to do so with the mechanisms that have served us well thus far.

Appendix 1. Courses currently offered that are relevant to AI/ML

I. The nuts and bolts of AI/ML, such as big data, software engineering/coding, database systems, and prompt engineering

A. Large Language Models: Technology and Applications (elective), Kyle Jensen (Entrepreneurship) and K. Sudhir (Marketing)
This course introduces students to the technology of large-language models (LLMs) such as ChatGPT and business applications thereof. Students will write substantial amounts of code individually in Python.

B. Big Data (elective), Vahideh Manshadi (Operations)
Cheap storage and computing power have enabled the gathering and analysis of an unprecedented amount of data on everything from genetic health-risk profiles to real-time Wall Street diaper consumption. To take advantage of these massive datasets, new statistical tools and ideas have been developed, and this body of knowledge is sometimes referred to as *data science*. The aim of this course is to provide a gentle tour of the business and industry applications of data science. Through the examples we will study, you will gain an intuitive understanding of the underlying data-analytic techniques that are often applicable to a wider class of problems. After completing this course you will have developed an appreciation for what opportunities exist for use of data within your organization.

II. Specific applications for AI/ML within organizations and fields of study (asset pricing, predictive decision making, social media analytics, etc.)

A. Financial Econometrics and Machine Learning (elective), Bryan T. Kelly (Finance)
Empirical work is the foundation of great economics. Theory is also the foundation of great economics. Theorists work on closing the gap between theory and reality. Empiricists are explorers who map the unchartered territory between theory and reality. This is a division-of-labor view of Popper's philosophy of science. This course is designed to help build a skill set for pushing the empirical side of this proposition that is particularly tailored to asset pricing research. (There is also a PhD version of the course.)

B. Big Data & Customer Analytics (elective), Kosuke Uetake (Marketing)
In today's information economy, companies have access to data about markets, products, customers, and much more. When deciding on strategic issues, such as pricing and advertising, targeting these data can be very valuable to companies if used correctly. This course will provide you with the tools and methods that will allow you to leverage data to help shape a marketing strategy from a quantitative perspective.

C. Social Media Analytics (elective), Tauhid Zaman (Operations)
This course will expose students to the key quantitative tools needed to analyze and create social media data. Topics include measuring social media sentiment, AB testing content

engagement, finding influencers, segmenting users, visualizing network data, creating persuasive text- and image-based content, and running automated influence campaigns. We will use a variety of AI and statistical tools in this course. These include analysis tools, such as transformer neural networks and clustering algorithms along with generative tools, such as ChatGPT (text generation) and MidJourney (image generation).

III. Macro-level phenomena related to AI's impact on markets, competition, regulation, the global workforce, and society

A. AI Strategy & Marketing (elective), Vineet Kumar (Marketing)
AI is a general-purpose technology that has the potential to transform many aspects of business and society. In business, the impact ranges from commonplace predictive improvements at one end of the spectrum to opportunities for creating entirely new markets at the other. As background, the course will briefly introduce students to AI/ML methods comprising Unsupervised, Supervised and Reinforcement Learning. Through a combination of lectures and case studies, we will evaluate how to integrate AI into decision making and examine the strategic choices facing companies developing and using AI/ML technologies. We will evaluate how both consumers and decision makers evaluate decisions made by AI systems and the feasibility of explainable AI. The course will also examine issues at the intersection of AI and society including fairness and bias, which are proving to be especially challenging, and an understanding of how both consumers and decision makers evaluate decisions made by AI systems will be developed.

B. Designing & Leading Organizations (MBA for Executives), Balázs Kovács (Organizational Behavior)
Organizations aggregate individual efforts toward a goal. They may take multiple shapes and forms to coordinate and motivate individuals. In this course, we overview the major forms in which organizations are designed. Besides analyzing the classic forms of organization design, this course puts an emphasis on novel opportunities and challenges that have emerged due to recent processes, such as globalization, network economies, the internet, big data, or crowdsourcing.

C. Session 4, "Changes in What Work Is Being Done," from *The Workforce* (core course), Laura Adler (Organizational Behavior)
Recent years have seen the rise of powerful new technologies including machine learning and AI. These tools raise important ethical questions about privacy, discrimination, and the relationship between humans and machines. These new technologies also raise urgent questions about the future of work: As we embrace the automation of nonroutine tasks, what will happen to people's jobs? In this session, we will discuss the fundamentals of AI and its ethics with a focus on the impact of AI on jobs. Drawing on lessons from historical cases of automation, we will consider how workers can be retrained for the kinds of jobs that will be needed in the future, including those that complement AI and jobs in areas, such as healthcare, that resist automation.

D. Economic Analysis of High-Tech Industries (elective), Ted Snyder (Economics)
This course applies economic concepts from industrial organization (IO) to high-tech industries. We will focus on four industry verticals: (1) mobility (EVs, ride-sharing, space travel, etc.), (2) video (streaming, gaming, Meta, etc.), (3) eCommerce, and (4) payment systems. We will analyze these verticals across three regions: (1) China, (2) the EU, and (3) the U.S. Our analyses will account for major forces, such as AI, the advance to 5G, increased China–U.S. tensions, and more aggressive competition policies and regulation. Students will be assigned to teams that focus on one cell in the matrix of three regions and four verticals, e.g., payment systems in the EU. Along with the final team project, the course requirements include class participation, a quiz, individual projects, team projects, and providing feedback to classmates.

E. The Science of Experiences and Well-Being (elective), Gal Zauberman (Marketing)
The goal of this course is to provide an in-depth exploration of the role of experiences in business and people's lives. Experiences play a vital role driving overall well being, from momentary enjoyment to life satisfaction and sense of self. This course explores a wide range of questions surrounding experiences and well being based on current scientific evidence. By developing an evidence-based and nuanced understanding of these issues, this course will aid you in designing better experiences for employees, for customers, and for yourself, allowing a more effective management of well being.

F. Build a Metaverse Strategy (elective), Brett Prescott
The metaverse is dynamic and has the potential to shape the future of human interaction. Citi analysts estimate that the metaverse will have a total addressable market of up to $13 trillion by 2030 with 5 billion users. Build a Metaverse Strategy approaches the topic through the lens of marketing, innovation, and commerce — ultimately identifying how businesses might leverage the metaverse to drive profitable growth. This course will immerse students in all major metaverse and web3 technology platforms. Students will leverage academic frameworks to evaluate emerging web3 technologies, identify business opportunities and propose a long-term metaverse business strategy for a Fortune 100 company. Through a nontechnical lens, students will learn how to build B2B and B2C marketing strategies that incorporate web3 technologies, including: nonfungible tokens (NFTs), the blockchain, smart contracts, the creator economy, avatars, decentralized data, decentralized finance (DeFI), decentralized autonomous organizations (DAOs), AI, augmented reality (AR), virtual reality (VR), and extended reality (XR). Academic frameworks that will be explored and applied include customer value proposition (CVP), profit formula, key resources, key processes, stage gate process, sizing up new marketplaces using total addressable market (TAM), serviceable available market (SAM), serviceable obtainable market (SOM), technology diffusion, and defining brand relevance using the brand vision model.

Appendix 2: Survey Instrument

Q1 Please select an answer for each of the following prompts:

	Never	Limited	Somewhat	Significant Degree	Core to What I Do	Not Now but in the Future
I use Artificial AI/ML methods in my research	○	○	○	○	○	○
In my research, I develop new AI/ML methodologies	○	○	○	○	○	○
In my research, I use the cloud to conduct my AI/ML work	○	○	○	○	○	○
In my course(s), I teach how to use AI/ML methods	○	○	○	○	○	○
In my course(s), I teach about how AI/ML is impacting the nature of work	○	○	○	○	○	○
In my course(s), I teach about how AI/ML is impacting markets, industries, and/or society	○	○	○	○	○	○

Q2 I find the Yale infrastructure (servers, laptops, etc.) sufficient for my AI/ML work:

○ Strongly Agree

○ Agree

○ Neither Agree nor Disagree

○ Disagree

○ Strongly Disagree

Q3 Describe the tools and infrastructure you use when conducting research involving AI/ML. (For example, are you using central Yale computing resources or are you using cloud-based services external to Yale? What kind of technical support do you receive? Is it from SOM or Yale staff, research assistants you hire, etc.)

Q4 Describe the tools, infrastructure, budget, and/or training that you think would significantly improve your ability to work with AI/ML in your research and/or teaching, but is not currently available to you.

Q5 Briefly describe your future plans for using AI/ML in research and teaching to the degree that they significantly differ from what you are doing now.

Q6 We would like to be able to follow up on your response based on the needs and experiences you've identified. To that end, please provide your name in the field below if you are comfortable doing that.

Appendix 3: Conferences, Seminars, and Speakers

Yale SOM's International Center for Finance in coordination with the Yale SOM AI Group hosted a multidisciplinary Responsible AI in Global Business Conference in March 2024, with the goal of collaboratively navigating the path of responsible AI development and adoption.

Several seminars have been offered at Yale SOM in the past year that focus on AI, including a talk by Professor Soroush Saghafian of the Harvard Kennedy School entitled "Making AI Impactful in Healthcare" and a talk by Professor Amy Ward of the Booth School of Business entitled "When Machine Learning Impacts Resource Allocation Decisions: OM in Criminal Justice Systems."

Additionally, many less-formal events with themes related to AI have been organized by SOM faculty and clubs. A sampling follows:

- Professor Kyle Jensen (SOM's Program in Entrepreneurship) hosted a series on Effective Use of AI to describe the conceptual underpinnings of modern AI tools, such as ChatGPT and Midjourney.
- Candace Harris, CEO and founder of Myavana, spoke about her company and how it uses AI to cater to its clients' hair needs.
- Dr. Casey King (PhD '10) lectured on business in the AI era to inform and equip the SOM community with the mindset and tools needed to compete in our ever-changing technological world.
- Yale's Professional Communication Center organized a discussion for SOM students on possibilities, limitations, and ethical considerations of using AI writing tools, such as ChatGPT.
- Yale Professor Brian Scassellati shared "ten things to know about AI" to help the SOM community be better equipped to understand relevant AI topics.
- A Q&A with Jen Hollingsworth, chief commercial officer at Flawless, was hosted to discuss how generative AI is revolutionizing filmmaking.

Yale School of Medicine

Report to the Yale Task Force on Artificial Intelligence: Basic Science and Clinical Research Panels

APRIL 1, 2024

Introduction

In response to the provost's call for faculty-led panels to brainstorm on school-specific approaches to develop and harness new artificial intelligence (AI) technologies, Yale School of Medicine (YSM) convened panels in the basic sciences and clinical research. Deputy Deans Tony Koleske (Basic Science Research), Brian Smith (Clinical and Translational Research), Lucila Ohno-Machado (Biomedical Informatics), and Peggy McGovern (Clinical Affairs) conferred with Dean Nancy Brown to identify faculty who bring unique expertise and experience in these areas that can inform school-wide strategic planning around AI.

Yale Medicine CEO and Deputy Dean for Clinical Affairs McGovern and Yale Health CEO Jason Fish convened a university-wide panel on the use of AI in clinical practice. In addition, Assistant Dean for Education Jaideep Talwalkar participated in a university-wide education panel.

This report focuses on the deliberations and presentations of the Basic Science and Clinical Research Panels. Recommended university-wide initiatives and investments are presented in Appendix 1. Membership of the YSM Basic Science and Clinical Research panels along with questions provided to facilitate discussions are presented in Appendix 2.

Key Findings

Conversations among the Basic Science and Clinical Research Panels revealed many common themes:

- Yale is leading in several areas, particularly clinical research in AI.
- The ability to collaborate and bring together expertise across the university offers a distinct advantage.
- We must find new ways to collaborate with industry, government, and academic partners. The domain expertise of our faculty is highly valued by industry and government leaders in AI who have access to large amounts of data but do not have the expertise to inform training and implementation of models.
- There is a growing shortage of graduate students, postdoctoral fellows, and research-rank faculty trained in AI at Yale University and in general.
- There is a shortage of computing power, particularly to handle sensitive data, at Yale University and Yale School of Medicine.

- YSM faculty have been creative in attracting undergraduates and others to their laboratories and in leveraging other institutions to increase access to computational resources.

Existing Strengths and Differentiators

Collaborative AI methods development and applications in basic science

YSM and its university partners have a unique opportunity to lead AI in science and discovery based upon our data, domain expertise, and computational expertise. Our basic science researchers are generating more high-throughput, multimodal data than ever before; our domain expertise within the life sciences is high; and we are benefitting from growing computational expertise, both within the school and across the university in the departments of computer science (CS), biomedical informatics and data science (BIDS), and biostatistics as well as in institutes and centers, such as the Institute for Foundations of Data Science and the Schmidt Program on Artificial Intelligence, Emerging Technologies, and National Power.

This combination of expertise enables us to create AI models that both incorporate vast existing knowledge in biology and are trained on biomedically relevant tasks. The primary goal of these new AI models is to generate causal (mechanistic) relationships and new hypotheses for further experimental follow-up. A key advantage to developing these models at Yale is that we have the capability to support end-to-end (translational) research through methods development, validation, initial testing, and active learning.

Among the many areas of existing, multidisciplinary strength and synergistic opportunity that facilitate the application of AI in the analysis of large data are:

- Immunology and infectious disease — Center for Systems Engineering and Immunology (CSEI), Yale Center for Infection and Immunity (YCII), Human and Translational Immunology (HTI), Chan Zuckerberg Institute (CZI Biohub)
- Genome interpretation and personalized medicine — Yale Center for Genomic Analysis (YCGA), Yale Center for Genomic Health (YCGH), the Department of Genetics
- Neuroscience — the Department of Neuroscience, Kavli Institute, Wu Tsai Institute (WTI), the Department of Cell Biology, Biomedical Imaging
- Cancer — Pathology, Yale Cancer Center, Cancer Biology Institute
- Stem and developmental cell biology — the Department of Cell Biology, Stem Cell Center

A shared electronic health record across diverse populations of patients

Yale Medicine (YM) and Yale New Haven Health System (YNHHS) have shared an electronic health record since 2013 and jointly employ a chief information officer and, now, a chief research information officer. Data from our clinical-data warehouse are routinely loaded into a computational health platform and augmented with real-time data feeds from multiple data sources, including physiologic monitoring systems, imaging, clinical integration engines, and third-party tools for data enrichment. The diversity of our patient population enables us to develop

and test AI technologies to ensure generalizability and freedom from bias. YNHHS and YSM have recently contracted to join COSMOS in making data from other national EPIC users accessible. Faculty are also participating investigators in the All of Us and Million Veterans research programs.

Significant investments in biomedical informatics and data science

Over the last several years YSM has made substantial investments in data science. Following up on a 2015 task force recommendation, in 2021 the school launched a national search for the inaugural leader of a new free-standing section of biomedical informatics and data science. This section now includes twelve ladder-track faculty, thirteen research-rank faculty, and forty secondary faculty. YSM faculty bring expertise in every facet of AI applicable to clinical research (and healthcare), including natural language processing (NLP), generative AI, predictive modeling, and fairness/ethical assessment.

Significant investment in the creation of a central biorepository

While many individual investigators had developed robust disease- and tissue-specific repositories, the coordination of these repositories, electronic health records, omics data, and imaging is necessary to train and make large integrated models accessible. To this end, YSM created a centralized Biobank approximately two years ago. This is a coordinated biorepository that allows for the alignment of collected tissue, plasma, serum and genetic samples, and clinical data using common processes and information systems. The goal is to streamline the distribution of specimens to investigators and promote collaboration of partners across the university.

Broad experience in AI for clinical research

Clinical research encompasses discovery across the spectrum by predicting the effectiveness of novel discoveries, translating discoveries to patient care, evaluating the effects of implementation on outcomes, and building community partnerships. With over fifty YSM (SEAS and WTI)-YNHHS projects related to AI or predictive modeling across virtually all clinical disciplines, Yale is already leading in AI clinical research. The panel presented several examples.

One is AI-enhanced screening of heart disease from wearables. Among adults, 5 percent suffer from cardiac muscle disorders with major implications for outcomes, but most are diagnosed too late, as the diagnosis requires advanced cardiac imaging. Rohan Khera, MD, MS, and his team tested the hypothesis that patterns of electrical activity of the heart that can be captured on ECGs, and now on many wearable devices, have signatures of structural disorders. Human readers cannot detect these disease signatures, but AI can be used to enable diagnostic inference from this data. In February 2024, the U.S. Food and Drug Administration (FDA) granted breakthrough device designation for the team's ECGvision-TTR© technology for early detection of a genetic form of cardiomyopathy.

YSM faculty Hua Xu and Annie Hartley have developed and currently maintain some of the world's best-performing open-source, open-access large-language models (LLMs) for medicine, including both generalist and specialty-specific models. In the hands of clinicians, these models can catch errors, improve diagnosis, and in some instances (i.e., psychotherapy), even deliver treatment.

Significantly, LLMs can also be used to expand access to care to populations that are typically unreachable. To assess the benefit and potential harms of LLMs, YSM faculty have launched MOOVE, a massive online open validation and evaluation platform that aims to leverage the expert community at Yale while formalizing a community-driven, continuous real-world alignment with our profession's ethical standards.

Sanjay Aneja demonstrated AI algorithms that augment molecular biomarkers and enhance the ability to predict survival from lung cancer based on imaging studies. Colleagues in radiology and biomedical imaging, such as Melissa Davis, serve as national leaders.

Degree-granting and certificate programs in informatics and AI

Educational programs led and taught by biomedical informatics and data science (BIDS) faculty include an MS in health informatics; an MHS with a clinical informatics and data-science track; and an MS in computational biology and bioinformatics.

In 2024, Professor Xenophon Papademetris launched an online certificate program, Medical Software and Medical Artificial Intelligence, which provides four-week modules on the Introduction to Medical Software, Introduction to Artificial Intelligence, Medical Software with AI, and Current and Emerging Applications of AI in Medicine.

The Yale–Bohringer Ingelheim Data Science Fellowship Program, led by the Center for Biomedical Data Science, offers a synergistic model for training scientists in AI who can cross industry-academia boundaries.

Vision and Potential for Leadership

Based on the reflections of the YSM Basic Science and Clinical Research Panels, we have formulated a draft of goals for YSM. All of these require cross-campus collaboration. School goals and recommendations for university-wide, cross-cutting investments (Appendix 1) are closely related.

YSM Goals

- Create a collaborative AI research ecosystem in the basic sciences across the university. Continue to integrate data from the electronic health record, omics sources, and advanced phenotyping (e.g., imaging) to create a shared resource that minimizes the risks associated with data sharing and enables AI workflows.
- Use this AI research ecosystem across the university to develop and disseminate standards for data quality and methodological validation. Yale University, YSM, and academia in general should establish standards for data cleaning, rigor, and reproducibility.
- Leverage the Harvey Cushing/John Hay Whitney Medical Library and the broader university library system to increase access to external databases through institutional subscriptions. Utilize the efficiency of this central resource to make more AI tools and models available to faculty.
- Partner with legal, ethical, economic, and policy experts across the university and community

groups to set standards for the ethical implementation of AI in clinical research and practice.

- Participate in rigorous processes and policies for evaluating, approving, and monitoring healthcare-AI products prior to implementation in clinical practice in YNHHS and YM to identify areas that require additional research.
- In collaboration with YNHHS and Yale Center for Research Computing (YCRC), establish a robust HIPAA-compliant infrastructure to support the AI lifecycle, thus ensuring a secure compute environment that is accessible to users across the university.
- Develop a strategy and service for independent validation of AI models in diverse and underserved populations. Such a service would attract collaborations with industry, government, and nongovernment partners.
- Create outreach and education forums about AI for community groups.
- Increase AI literacy in basic statistics and machine learning by providing practical AI training for all. This could be offered through a variety of mechanisms, such as hands-on "bootcamps" for incoming students (e.g., neuro-quantitative and genomics), practical training courses and workshops, and expansion of in-person and online consultations available through YCRC, Harvey Cushing/John Hay Whitney Medical Library, and StatLab.
- Continue to develop creative approaches to recruiting AI postdoctoral trainees: interdepartmental, coordinated recruitment efforts; partnerships with industry (e.g. Boehringer Ingelheim); and sponsored co-mentorship programs (e.g., WTI fellows, Yale Institute for Foundations of Data Science).
- Review the undergraduate and graduate medical curricula to introduce concepts related to critical evaluation of AI, etc.

Appendix 1: Recommended University-wide Initiatives and Investments

- Increase computational power through acquisition of graphics processing unit (GPU) clusters and cloud time as well as through bidirectional partnerships with other academic institutions, industry, and governmental and nongovernmental organizations.
- In collaboration with YNHHS and YSM, establish a robust HIPAA-compliant infrastructure to support the AI lifecycle, thus ensuring a secure computing environment that is accessible to university-wide users.
- Leverage the convening power of the Provost's Office to create spaces for cross-disciplinary collaboration in AI.
- Increase the number of computational graduate students. This should include increasing the number of students in computational biology and bioinformatics (CBB), but it should also include cross-training graduate students in the biological sciences in computation and students in the computational sciences in biological science. It is recommended that biological and biomedical sciences (BBS) conduct a curriculum review and consider integrating teaching in AI across domain tracks.

- Consider opportunities for cross courses or dual majors in computer sciences and common premedical majors. Yale College should consider AI courses for premedical students.
- Codify standards for scholarly contributions to multidisciplinary work in AI in appointment and promotions policies.
- Develop academic homes for those engaged in ethics related to AI.
- Enhance expertise in AI within Yale Ventures.

Appendix 2

Basic Science Panel Members

- Kristen Brennand, PhD, Elizabeth Mears and House Jameson Professor of Psychiatry; Codirector, Science Fellows Program
- Hyunghoon Cho, PhD, Assistant Professor of Biomedical Informatics and Data Science
- Peter Gershkovich, MD, MHA, Associate Professor of Pathology; Director, Section of Pathology Informatics and Cancer Data Science, Pathology
- Ira Hall, PhD, Professor of Genetics, Director of the Yale Center for Genomic Health
- Steven Kleinstein, PhD, Anthony N. Brady Professor of Pathology; Codirector of Graduate Studies, Computational Biology and Bioinformatics
- Yuval Kluger, PhD, Anthony N. Brady Professor of Pathology
- Smita Krishnaswamy, PhD, Associate Professor of Genetics and of Computer Science
- Zeynep Erson Omay, PhD, Assistant Professor, Neurosurgery, Biomedical Informatics and Data Science
- Xenophon Papademetris, PhD, Professor of Biomedical Informatics & Data Science, and Radiology & Biomedical Imaging; Director of Image Processing and Analysis, Bioimaging Sciences, Radiology and Biomedical Imaging
- Steven Reilly, PhD, Assistant Professor of Genetics

Basic Science Questions

- How is AI being used currently in basic science within YSM and by collaborators across the university?
- Are there areas of strength either in AI or in other technologies relevant to AI (e.g., genomics, single-cell RNA sequencing, imaging) that differentiate us from our peers?
- Where are the gaps?
- What university policies might facilitate collaboration and advances in AI in basic science?
- What educational resources are needed to train experts, users, graduate students, principal investigators, etc.?
- What collaborations across institutions (universities, industry) would facilitate AI development?

- What is needed from Yale Ventures and IP/copyright policy?
- How should discoveries made through AI be validated? How do we ensure rigor?

Clinical Research Panel Members

- Sanjay Aneja, MD, Assistant Professor of Therapeutic Radiology; Director of Clinical Informatics, Therapeutic Radiology; Director Medical School Clerkship, Therapeutic Radiology; Medical School Thesis Oversight, Therapeutic Radiology; Radiation Safety, Therapeutic Radiology; Assistant Cancer Center Director, Bioinformatics
- Melissa Davis, MD, MBA, Associate Professor of Radiology and Biomedical Imaging; Vice Chair for Imaging Informatics, Radiology & Biomedical Imaging
- Mary-Anne "Annie" Hartley, MD, PhD, MPH, Assistant Professor of Biomedical Informatics and Data Science; Affiliated Faculty, Yale Institute for Global Health
- Rohan Khera, MD, MS, Assistant Professor of Medicine (Cardiovascular Medicine) and of Biostatistics (Health Informatics); Clinical Director, Center for Health Informatics and Analytics, YNHH/Yale Center for Outcomes Research & Evaluation (CORE); Director, Cardiovascular Data Science Lab (CarDS)
- Daniella Meeker, PhD, Associate Professor of Biomedical Informatics & Data Science; Chief Research Information Officer, Yale School of Medicine and Yale New Haven Health System
- Marcella Nunez-Smith, MD, MHS, Associate Dean for Health Equity Research and C.N.H. Long Professor of Internal Medicine (General Medicine), of Epidemiology (Chronic Disease), and of Public Health (Social And Behavioral Sciences) and Professor of Internal Medicine (General Medicine); Affiliated Faculty, Yale Institute for Global Health; Founding Director, Equity Research and Innovation Center (ERIC), Yale School of Medicine; Director, Center for Research Engagement (CRE); Director, Center for Community Engagement and Health Equity; Deputy Director for Health Equity Research and Workforce Development, Yale Center for Clinical Investigation (YCCI); Director, Pozen-Commonwealth Fund Fellowship in Health Equity Leadership
- Wade Schulz, MD, PhD, Assistant Professor; Director of Informatics, Laboratory Medicine; Director, CORE Center for Computational Health, Center for Outcomes Research & Evaluation (CORE)
- Hua Xu, PhD, Robert T. McCluskey Professor of Biomedical Informatics and Data Science; Vice Chair for Research and Development, Section of Biomedical Informatics and Data Science; Assistant Dean for Biomedical Informatics, Yale School of Medicine

Clinical Research Questions

- What is currently going on in AI in clinical research in YSM, YNHHS, the broader university and community? What would you highlight?
- What are the opportunities for AI in clinical research in the future, thinking broadly, beyond the walls of one institution?

- Are there areas that differentiate us from others? Where are the gaps?
- What are the cross-cutting areas in which the university should invest?
- What are the educational needs for investigators or for clinicians interpreting research? How should we train students, existing investigators, and clinicians?
- What are the risks of AI in clinical research? How do we mitigate those risks?
- What is the impact of AI on inclusion in clinical research? What is the risk of bias?
- How do we ensure that we protect privacy while using AI methods?
- How do we need to validate findings from AI studies? Are there learnings from FDA decisions regarding new AI tools? Who should set the standards?
- How do we need to educate policy makers?

Yale School of Music AI Summary

MARCH 29, 2024

The School of Music's mission currently centers on the performance and composition of notated music in and around the Euro-diasporic tradition often referred to as "classical music," primarily for acoustic instruments and voices. Students are trained in the art of interpreting this music in the act of live performance, informed by analysis of its structure and its relationships with complex historical, stylistic, and intercultural precedents stretching over several centuries. Reflections on the ways in which the school can take a leadership role in the development and use of artificial intelligence (AI) must take into account these current practices. As a school of music within a top-caliber research university, we are well positioned to confront the limitations of AI and shape its practical applications. Just as important, we are also poised to address its social and socio-economic impact on classical music and beyond.

Given the dominance of visual culture in the internet age and the small market share of classical music within our larger cultural milieu, it is no surprise that this music has not been a primary focus of AI innovation to date. At present, machine-learning models are severely limited in their capacity to interpret sonic data and to analyze musical scores and music-historical sources. While existing software can convert scores into sound and vice versa under certain controlled circumstances, current AI fails when asked for the kinds of information, whether analytic or historical, that could stimulate convincing live performances. This is due both to the specialized vocabulary characteristic of scholarly writing about music and to the limited corpus of that writing. Given the data used to develop large-language models, these models produce plausible-sounding yet often faulty and unreliable discourse on music, which means they are generally incapable of helping performers in meaningful ways with the interpretive choices that define their artistry. The limitation of current AI models extends to the classroom, where most of our time with AI presently involves helping students understand its musical incapacities rather than using it to empower their musical creativity.

Despite these fundamental limitations, however, the arenas of scholarship, performance, and composition provide concrete opportunities for AI exploration at the School of Music. In the scholarly arena, AI software already exists that can transcribe language in historical sources into modern type. Such software can be as helpful in music studies as in any other humanistic or historical field. Most excitingly for the school, this technology provides a valuable tool to transcribe and analyze written texts from marginalized communities, which means it can support efforts to broaden our musical horizons beyond the boundaries of the traditional European canon while also enriching the canon itself via dialogue with musics of other cultures. Exploration in this direction can involve collaboration with the Department of Music and the Institute of Sacred Music.

In the performance arena, AI has the potential to provide adaptive accompaniment for solo musicians in rehearsal settings, with machine learning adjusting the tempo of digitally encoded

music or prerecorded sound in real time to match the playing tendencies and decisions of the soloist. As of yet, this software does not approach a plausible substitute for human music making in actual performance, but with sufficient advances along these lines, more efficient modes of rehearsal will become more widespread, and musicians otherwise unable to perform certain repertoire will find opportunities to do so. Potential beneficiaries of such applications extend beyond the school's own students and faculty to include performers from across the university, public school students served by our Music in Schools program, and ultimately anyone whom the university wishes to reach via digital means. A natural hub for further research in these directions is the school's Center for Studies in Music Technology.

The arena of composition currently offers perhaps the most promising opportunities for AI-based innovation at the school, along with some of the most troubling implications. It is our consensus that machine learning will very soon be able to produce music in a variety of styles at a level commensurate with commercial uses: background music, mood music, underscoring in film, video game music, and many other popular genres. Indeed, AI is already reshaping the ecosystems of commercial music production, both by creating music that would otherwise be made by musicians and by fine tuning the streaming algorithms that shape the musical taste and experience of the vast majority of listeners. The socio-economic results are not limited to popular entertainment, but also affect classical musicians directly. Many composers of classical music support themselves financially via commercial composition, and some classically trained performers make substantial portions of their living via participation in popular genres. The elimination or reduction of live-human employment in commercial musics and the algorithmic homogenization of musical preferences threaten the foundations of the classical music ecosystem as well. All conservatories and schools of music will need to adjust their pedagogical emphases to address this reality.

At this point it is less clear how soon, if ever, AI will be able to simulate, for professional musicians, the act of composition in classical or "art music" settings, since those settings privilege methodological originality, stylistic multivalence, and carefully refined relationships between new music and existing repertoire. But even in classical settings, AI can play crucial roles, for good and ill. One obvious benefit of AI is its potential as a tool to generate or edit sound in electronic or multimedia compositions, and that potential should be pursued explicitly at the school. In so doing, however, we are positioning AI to be a critical mediator of the sonic datasets that composers of either electronic or acoustic music draw upon in the creation of new works. The danger we perceive here, as with AI in all creative arts, is its convergent tendency toward homogeneity, its regression toward the mean, which has the potential to limit the choices that humans fluent in AI-based models perceive to be possible in the first place. For this reason, alongside the development of new AI-based composition tools, another fruitful arena of AI leadership for the school lies in the curation of divergent sonic datasets and their foregrounding in the education of our students. Again, the Center for Studies in Music Technology is a natural hub for research and teaching in this arena.

The School of Music is in a position to make important and lasting contributions to Yale's conversations about AI-driven technology. Alongside the possible innovations detailed above in the fields of scholarship, performance, and composition, however, the school's most important contribution may be its focus on the human element that defines the performing arts. We believe that the coming hegemony of AI across a variety of cultural arenas will precipitate an urgent search for explicitly human ways of knowing and doing and that we will find them in live performances, be they in the concert hall, public schools, community centers, or the open air—wherever human bodies and minds are placed in dialogue with other human bodies and minds by the demands of a musical work and the unscripted impetus of the moment. In a future shaped by AI, live performance will be a crucial space for the renewal of the human and the working out of its relationship to the nonhuman.

Yale School of Nursing
Report to the Yale Task Force on Artificial Intelligence (YTAI)

APRIL 1, 2024

Introduction

Artificial intelligence (AI) holds significant potential across various domains in medical and nursing care and education. However, it also brings forth challenges that necessitate acknowledgment of and concerted efforts by institutions such as the Yale School of Nursing (YSN) to address them comprehensively across educational, research, and clinical arenas.

While AI capabilities have seen exponential growth in recent years, the notion of leveraging computer assistance to aggregate and analyze knowledge in medicine is not novel. The PubMed journal list alone now encompasses over 30,000 titles, and the sheer volume of new research published annually, even within specialized areas, has reached staggering proportions. Nevertheless, critical diagnostic or therapeutic insights may remain concealed within the vast expanse of literature. Enabling the discovery of such information represents one of the most promising applications of AI—a promise that is already being actualized.

Clarifying the relationship between AI and healthcare, particularly within the context of nursing education and practice, is imperative. YSN must navigate this landscape with precision, harnessing the potential benefits of AI while mitigating associated risks. Furthermore, in the clinical realm, AI can augment healthcare professionals' capabilities by facilitating data-driven decision making and streamlining processes. However, ethical considerations, including patient privacy, algorithmic bias, and accountability, demand meticulous attention.

The primary immediate concern with AI is its current state of immaturity, which often masks its true level of reliability. A common misconception holds that AI is infallible, yet at this stage of development it possesses significant fallibility, particularly in contexts crucial to nursing. In essence, AI has the capability to produce highly convincing but inaccurate results, akin to telling very convincing lies. A recent and egregious example of AI's potential to generate false information was observed in a legal filing submitted on behalf of attorney Michael Cohen, wherein legal citations provided were to cases that did not exist. In a medical context, such errors could have dire consequences, potentially leading to fatal outcomes rather than mere embarrassment. While apologizing for a mistake made in front of a judge might suffice, the prospect of having to apologize to a surviving spouse or other family member due to AI-generated errors underscores the gravity of the situation. In simpler terms, the ramifications of AI's inaccuracies in a nursing setting can be life-threatening, emphasizing the critical need for caution and vigilance in its implementation and interpretation.

As an initial step, and in response to the provost's call for faculty-led input, YSN conducted a faculty survey and convened a faculty panel on February 14, 2024, at the Poorvu Center for Teaching and Learning. The panel consisted of Dr. Christine Rodriquez, Dr. Zhao Ni, and Dr. S. Raquel Ramos.

Key Findings

- As a school of nursing, we encounter distinct challenges that extend beyond the typical concerns associated with the integration of AI. Our domain is governed by ethical and legal constraints and mandates that may not apply to other academic disciplines. Addressing these considerations often necessitates tailored approaches and accommodations specific to the operations of the YSN. This is particularly important because we are responsible for training future clinicians, who must navigate these ethical and legal complexities in their practice.
- We share with all other academic disciplines the need to understand the optimal ways in which students can use AI to enhance their education and the ways in which it can be misused in both the academic and clinical environments.
- Boundaries on deployment of AI by faculty and students need to be established to maintain ethical and educational objectives and to remain within already established legal limits.
- AI has enormous prospects for enhancing our ability to provide quicker, more accurate diagnosis and treatment.
- There will be unparalleled demand and competition for faculty and staff who are knowledgeable about AI and prepared to fully leverage its capabilities. As a school of nursing, we in turn must be prepared to educate students, faculty, and our clinical partners in effective use of AI.
- AI can provide highly individualized educational assistance and on-demand access to such assistance.
- Clinical use presents very substantial issues of patient privacy, data protection, and records accuracy.
- The healthcare professions as a whole need to decide the degree to which anonymized data can be aggregated and shared. This is the foundation on which AI is built, but it also presents obvious issues relating to privacy and security. This is further complicated by relationships that may be forged with commercial ventures.
- YSN is a global pioneer in many aspects of medically focused AI, including studying how to leverage AI chatbots to increase access to quality healthcare for vulnerable populations at high risk for HIV, harnessing machine learning to create algorithms for measuring sleep metrics, expanding variable capture in large databases to identify falls from radiography reports, and deployment of virtual-reality (VR) based high-fidelity manikin simulation to enhance nursing education.

Our Vision for AI's Future in Nursing

Currently, more than 90 percent of the YSN research faculty utilize some aspect of artificial intelligence in their work. To a lesser extent, AI is used for teaching/learning and clinical practice. Faculty members report employing AI in critical areas, such as data analysis and management,

employing machine learning for preventive interventions using electronic medical records, and developing avatars and AI responses to facilitate initial triage of patient needs.

AI is deeply embedded in the future of healthcare in this country, from education to clinical care to research. There is no question of *if* AI will be used; it is only a question of *how*.

Thoughtfully deployed and employed, AI has the potential to be truly transformative across the entire healthcare spectrum. It will be a powerful tool for improving the way we educate, the way we deliver clinical care, and the way we do research.

The YSN will continue being a commanding leader in the integration of AI in healthcare. We have the capability, commitment, and knowledge to define for our nursing colleagues how best to leverage the capabilities of AI, while also safeguarding the privacy of patients and the integrity of research efforts.

With AI, nursing will be better able to share educational and clinical knowledge on a global basis. "Classrooms" and "clinics" will be wherever there is an internet connection, and disparities in care will be reduced as answers and clinical-decision support become AI driven.

AI creates new options and opportunities for cooperation and collaboration on a national and international scale. New alliances will be formed inter-professionally and intra-professionally.

All of this will hasten the evolution of our profession from one for which the emphasis is on mechanical skills and knowledge to one where nurses make full use of their knowledge to understand and respond to the clinical needs of their patients and to better understand the implications of the research they do.

Nursing Education

We share with all other academic disciplines the need to understand the optimal ways in which students can use AI to enhance their education and the ways in which it can be misused in both academic and clinical environments.

In the educational setting, AI has a tremendous amount to offer.

The integration of AI into nursing education offers transformative opportunities, particularly in the realm of simulation-based learning.

- AI-driven advancements significantly enhance nursing simulation scenarios, which are crucial for educating a larger student body amidst declining availability of clinical placements. With sophisticated manikins and AI technology, students can gain competency in diverse and complex clinical situations not encountered during traditional rotations.
  - **AI-based avatars** can simulate patient interactions, while point-of-care coaches assist nurses in skill consolidation, emphasizing person-centered learning.
  - **Personalized learning algorithms** analyze student performance, identify areas needing improvement, and tailor learning materials and assessments to optimize overall success.

    - **Immersive-learning environments:** AI, in conjunction with extended reality (XR) technology, creates immersive environments that expand experiential learning within safe, controlled settings. The fidelity of these simulations ensures realistic replication of patient scenarios, enriching students' learning experiences.
    - **Innovative AI tools:** Introduction of novel AI tools allows the creation of dynamic patient characters with unique backstories, enabling real-time interactions with clinicians based on scenario details. Advancements in AI cinematography further enhance the construction of lifelike clinical scenarios, enriching hands-on experiences and clinical decision making. The overarching goal is to provide students with opportunities for hands-on experience, competency demonstration, and improved clinical decision making, all while ensuring the safety of human patients. Through these AI-enabled approaches, nursing education can evolve to meet the demands of modern healthcare delivery and prepare competent and compassionate nurses for the future.
- **On-demand assistance.** AI will provide immediate, comprehensive feedback as students engage with assignments. It will provide not only answers, but explanations and citations that encourage a deeper learning experience rather than mere memorization. Availability will be not only instantaneous, but also available 24 × 7 × 365.
- **Individualized tutoring**
    - Faculty will have the ability to design course materials that prompt student responses, with AI supplementing additional content based on those responses. This approach strengthens areas of knowledge weakness and guides students toward competency within segments defined by faculty.
    - By utilizing metrics, we can customize personalized learning experiences and establish AI-based tutoring systems. This is a future avenue for impactful learning accessibility, particularly for learners in rural and historically marginalized communities. Additionally, this fosters opportunities for remote learning and distance education through the integration of such technologies.
- **Currency:** AI will address the challenge of keeping learning materials current when medical knowledge evolves daily and even hourly. A summary in response to a question such as What are the reported adverse effects for Drug A? may be different on Wednesday than it was on Tuesday, and AI can keep pace with and incorporate new knowledge the moment it is published.
- **Multifaceted perspectives**
    - When utilized effectively, AI can provide a comprehensive overview of areas of nursing knowledge where opinions and data interpretations differ. This prompts students to consider various perspectives and justify their choices. Through AI, we can harness powerful algorithms to monitor and analyze a wide range of healthcare literature and clinical practice guidelines (CPGs).
    - Furthermore, AI holds promise in adaptive curriculum design (ACD) within higher education. This entails dynamically adjusting curriculum content and instructional materials. Imagine our current learning management systems (LMSs) powered by AI, ensuring relevance and currency

in emerging trends and innovations. This advancement is crucial for both early career and seasoned educators, as it alleviates faculty workload while meeting the evolving needs of diverse communities.

- Maximizing the utility and impact of AI for teaching will require teaching best practices for its use and awareness of its limitations. It may well be the case that "using AI in nursing" will need to be mandatory content threaded through all didactic and clinical education.

We also need to adjudicate the issue of where the boundaries are in terms of claiming authorship. It will be possible to have AI write a dissertation or a professional paper. How do we set limits on what an "author" must do to be the intellectual author of a dissertation, presentation, or paper.

Nursing Clinical Care

To comprehensively analyze the implications of integrating AI in nursing, we can examine four lenses: (1) ethical considerations, (2) privacy concerns, (3) regulatory compliance, and legal frameworks and (4) security risks. By approaching the topic from these perspectives, we can identify subthemes within each category.

Ethical considerations

- Biases and equity: Recognize that AI algorithms may perpetuate biases present in data, thus potentially exacerbating health disparities within disenfranchised and historically marginalized communities.
- Autonomy and accountability: Ensure transparency, accountability, and a commitment to beneficence and justice in AI-powered technologies; align them with accuracy and reliability through model validation and evaluation.

Privacy concerns

- Data safeguarding: Address the collection and safeguarding of collated data to protect patient privacy and confidentiality.
- Global standards: Expand considerations beyond HIPAA to include general data protection regulation (GDPR) standards, with potential utilization of informed consent.

Regulatory compliance and legal frameworks

- Adherence to frameworks: Ensure that AI applications comply with legal and regulatory frameworks, including discussions on liability and accountability for errors or adverse outcomes.

Security risks

- Cybersecurity threats: Mitigate risks of cyberattacks and breaches through encryption and threat-detection mechanisms to safeguard AI-powered systems and sensitive data.
- It is noteworthy that the White House Office of Science and Technology Policy has developed a Blueprint for an AI Bill of Rights in 2022. This includes provisions for safe and effective systems, addressing algorithmic discrimination, ensuring data privacy, providing notice and explanation,

and considering human alternatives and fallbacks. Incorporating this resource into our discussions on AI will help us stay informed about evolving themes and trends.

Nursing Research

It is in the research setting that AI has some of its greatest potential to stimulate and facilitate new ideas and provide new capabilities for understand large datasets.

- **Pattern recognition.** AI excels at mining large datasets for patterns that might otherwise be unseen by researchers manually assessing data.
- **Collaboration.** AI will bring to the fore information that helps forge partnerships with commercial and nonprofit entities — collaborations that will facilitate and accelerate research across a broad range of areas.
- **New datasets.** Because of its global reach and ubiquity, AI has the potential to enable researchers access to far larger datasets about a greater diversity of subjects than has previously been possible; they will also be able to find within those datasets significant similarities and differences.
- **Discovery.** With AI, researchers will have a tool that can lead to discovery of previously unseen/unknown relationships between, for example, a specific illness and specific medications.

YSN Goals: Infrastructure

YSN is committed to strategically leveraging AI to advance nursing education, research, and clinical practice. Our comprehensive action plan outlines specific initiatives aimed at harnessing the potential of AI while addressing ethical, privacy, and equity considerations:

Develop tailored training programs

- Create a structured curriculum and workshops focused on AI to equip future nurses with the necessary skills to effectively utilize this technology in clinical settings.
- Empower faculty with AI knowledge through workshops and personalized consultations to facilitate the integration of AI concepts into teaching and research endeavors.

Establish leadership in AI-supported education

- Position YSN as a national and global leader in AI-supported simulation and clinical education, leveraging innovative approaches to enhance learning experiences and patient outcomes.

Foster collaboration and resource optimization

- Engage in collaborative efforts to establish a centralized AI consulting resource serving all Yale healthcare schools, promoting efficiency and preventing duplication of efforts.

Promote awareness of ethical and privacy issues

- Facilitate ongoing discussions among students and faculty regarding ethical, privacy, and other critical issues pertaining to AI in healthcare.
- Evaluate and potentially revise Institutional Review Board (IRB) criteria to accommodate AI-related considerations, such as data releases and privacy waivers.

Address disparities and equity

- Consider the differential impact of AI on underserved and BIPOC communities when designing research and educational strategies, ensuring inclusivity and representation in decision-making processes.

Integrate ethical AI practices in research

- Continue to prioritize ethical considerations in groundbreaking research endeavors, emphasizing the responsible and transparent use of AI technologies.

YSN Goals: Research

Establish a research center on AI chatbots at YSN

- YSN aims to solidify its global leadership in AI by establishing a research center focusing on AI chatbots.
- Public sharing of AI models developed at YSN can further enhance its influence in the scientific community.
- Center would explore the integration of AI chatbots with augmented reality (AR) and virtual reality (VR) technologies.

  These endeavors will contribute to the development of innovative strategies and interventions for improving healthcare outcomes.

Exploration of robotics in healthcare

- Conduct formative research on the feasibility of integrating AI-based robots, such as Spot or Unitree B2, into healthcare systems.
- YSN's leadership in this field will lay the groundwork for future implementation science trials on leveraging robotics in healthcare.

Utilization of electronic health record data for research

- Support and encourage the use of electronic health record (EHR) data for research purposes, particularly in developing AI tools.
- Foster collaboration with departments of computer science to enhance YSN's participation in AI research teams.

Resources

Maximizing the impact of AI in nursing necessitates a collaborative effort between YSN, Yale University, and its various healthcare-related schools. Collaborative involvement may encompass:

Providing adequate computing resources

- Recognizing that AI tasks are computationally demanding, Yale will ensure the availability of sufficient computing resources to support AI initiatives.

Allocating sufficient funding for AI expertise

- Allocating resources to enable the recruitment of staff and faculty with expertise in AI and healthcare, thus ensuring a dedicated focus on advancing AI applications in the field of nursing.

Offering comprehensive university-wide AI technical support

- Yale will establish university-wide AI technical support services to assist faculty, staff, and students across all disciplines in leveraging AI technologies effectively.

Explore collaborations with vendors and technology companies to integrate emerging technologies, such as haptic feedback and 3-D printing, into healthcare training.

In tandem, the YSN will play a pivotal role by:

Defining and establishing internal AI expertise

- YSN will delineate specific roles for AI personnel within the school, facilitating the integration of AI expertise into various aspects of nursing education, research, and practice.

Creating a centralized AI infrastructure

- YSN will establish a centralized AI infrastructure tailored to the unique needs of nursing research and education, enabling the development of nurse-focused research projects powered by AI algorithms.

Developing innovative clinical education programs

- YSN will pioneer the development of cutting-edge clinical education programs that integrate AI technologies, ensuring that students are equipped with the skills and knowledge necessary to navigate AI-driven healthcare environments effectively.

Leveraging AI in simulation technologies

- YSN will fully leverage AI capabilities in manikin-based and other simulation technologies, offering students immersive learning experiences that compensate for the limitations in available clinical-placement opportunities.

By jointly addressing these areas, Yale University and the Yale School of Nursing can foster a conducive environment for the effective integration of AI in nursing, ultimately enhancing patient care outcomes and advancing the field of healthcare.

Yale School of Public Health
Report on Artificial Intelligence

MARCH 18, 2024

According to C.-E. A. Winslow, the founder of the Yale School of Public Health (YSPH), public health is "the science and art of preventing disease, prolonging life, and promoting health through the organized efforts and informed choices of society, organizations, public and private communities, and individuals." What these "organized efforts" and "informed choices" will look like in the future, however, is likely to be fundamentally different from what we have focused on over the last century. Climate change, mass migration, the transformation of the media ecosystem, and generational shifts in the workforce are fundamentally transforming the types of health threats that we face, the ways in which they spread, and best practices for monitoring and mitigating health risks on a planetary scale. The U.S. is at the bottom of health metrics, including life expectancy, among developed countries—despite spending a greater percentage of its gross domestic product on healthcare than any other nation; and historically marginalized populations in the U.S. consistently have worse health.

Meanwhile, health data—both traditional (such as that collected in clinical care, at pharmacies, and through blood work) and "trace" data (such as patterns of restaurant visitation, social media posts, and wearables) is accumulating quickly and is used without attention to privacy, bias, or accuracy.

In the face of these challenges, we at YSPH believe that it is necessary to re-imagine how we study and teach public health, with data as a foundation. To us, artificial intelligence (AI) serves as a catalyst for this inflection point: a tool for improving our use of increasingly complex data, a method to be improved upon, a facilitator of behavior change, and a tool to reduce instead of perpetrate health inequities.

As YSPH itself transitions into an independent school, we have the opportunity as well as the obligation to lead through the scientific and societal inflection points ahead. We are drawing on our school's existing expertise to transform the use and development of health-related AI in research, education, convenings, and practice. Some of this work uses AI to augment other health data and intervention methods, some develops new types of AI to reduce algorithmic uncertainty or improve representativeness, some focuses on creating guardrails for health safety, all combine our deep methodologic rigor and our foundation in health data with our humanistic roots and commitment to community as public health practitioners.

Specifically, we foresee—and are helping to create and implement—the benevolent potential for AI in enhancing the health of populations through equity-based approaches to data acquisition, algorithm development, and implementation. For example:

- Our faculty, staff, and students are collaborating with local community groups to use AI to improve how we gather data from traditional as well as nontraditional health sources, use these complex and

potentially messy data sources to identify "hot spots" of poor health, and then design and deliver appropriate resources to reduce the drivers of disease and injury in a way that enhances community investment and minimizes disparities.

- Our faculty, staff, and students are examining how AI can accelerate the speed of development of and improve the cultural relevance of health communication, ranging from chatbots to visual depictions; they are also using natural-language processing (NLP) and other types of AI to examine current communications (including social media posts).
- Our faculty, staff, and students are using AI to enhance the analysis of complex health-adjacent data sources (genomic, clinical, wearable, geospatial, and others) to identify the development of new health emergencies ranging from infectious epidemics to humanitarian crises.
- Our faculty and students are defining and then developing tools to reduce uncertainty in AI-powered decision making, particularly for rare diseases or marginalized and under-represented populations.

Similarly, we are actively working to research and scale AI-based tools that can identify and mitigate the negative health effects of AI, including bias, biosecurity, health information risks, and workforce disruption. For example:

- Our faculty, staff, and students are examining how AI may perpetuate health inequities and biases — whether through the type of data it uses or through its algorithmic outputs — and then are working to mitigate this bias, whether through new algorithm development, through calling attention to potential data disparities and under-represented populations, or through training the community workforce. (We held an international gathering on public health data equity in April 2024, to help accelerate.)
- Our faculty are leading national and international efforts to mitigate biosecurity risks created by AI, both on the policy and technical sides, working with the White House, national advisory groups, and international governance structures.
- Our faculty and staff are thinking about how AI will disrupt the public health workforce, developing trainings for community public health workers as well as incorporating AI into the classroom.
- Reflecting our and others' concern about how AI endangers the health of populations by facilitating the development and dissemination of disinformation, our faculty are participating in multiple national workgroups to create both scientific knowledge and policy recommendations, often in partnership with big tech.

As these examples illustrate, YSPH envisions the future of AI and health in a way in which AI serves as both a methodological innovation and a tool for re-imagining our field of public health, in collaboration with the communities we serve — and for fundamentally changing how we work with other fields, ranging from medicine to management to math, to enhance the health of populations.

We have collectively outlined three major areas of investment and focus, both within our own school (YSPH) and across the university (U). Internally, this work will be led by our new senior

associate dean of public health data science and data equity, who will join Yale in August 2024, as well as multiple existing faculty, staff, and students.

1. We must navigate dynamics on data availability, quality, completeness, governance, equity, privacy, and security, with a focus on enhancing trust in the data we collect and in our use of it.

    A. In partnership with community members, our development of methods for AI development and co-ownership is emergent. We must support community-participatory approaches to understand barriers to, then engage individuals in, data collection and analysis.
        i. **YSPH:** We will provide pilot awards to facilitate this work (see also 2b, below) and support faculty in submitting relevant grants.
        ii. **YSPH:** In collaboration with community members multiple YSPH researchers are considering how to enhance equity in data sources. We convened an international group of scholars, policymakers, and industry professionals in April 2024 to discuss this issue (in collaboration with the Macmillan Center, including colleagues across Yale), and will provide recommendations that arose from this meeting.
        iii. **YSPH + U:** Facilitation of interdisciplinary collaborations to improve trust, reliability, and representation of health data more substantively (in collaboration with the community) is needed.
    B. Although local electronic health-record data is helpful, what is more essential is the ability to access and synthesize federated datasets with other researchers across the globe. For example, the ability to use data from Cosmos (EPIC's national research compilation), 23andMe, the UK Biobank, the National Institutes of Health's All of Us Research Program, state-specific labor and health data, wearable-device data, and national satellite data on particulate matter and human migration, matter deeply.
        i. **YSPH:** We have begun an assessment of which large health datasets are being used by our faculty. In the years ahead, we plan to internally enhance the quality and quantity of dataset access, with global and domestic equity in mind.
        ii. **U:** There is no reason to duplicate data sources or to wall them off from our collaborators. We are enthusiastic to continue to collaborate with the Data-Intensive Social Science Center, the Biomedical Informatics & Data Science group, the Tobin Center for Economic Policy, and others to enhance our collaborative access to high-complexity data.
    C. YSPH uses a variety of tools and types of AI and works with a variety of providers of compute (including the Yale Center for Research Computing [YCRC] as well as state, national, and international collaborations and partners). Investments in core data-gathering and -storage enterprises are necessary.
        i. **YSPH:** We are considering how best to support increasing use of high-throughput, compute-intensive analytic techniques. This will represent a large financial investment on the part of the school.

ii. **YSPH + U:** Resources should be made available at the unit and university level for license fees for software and AI tools, respecting the diverse and specialized needs of the campus.
iii. **YSPH + U:** YSPH is committed to contracting and collaborating with researchers, governments, communities, and industry leaders across the globe. To develop these essential external and global partnerships—in data acquisition as well as in massive computation, data storage, and optimal workflow design—support is needed from the Office of the General Counsel, the Office of the Vice Provost for Research, Information Technology, and others.
iv. **U:** To support the development of large AI models (including, but not limited to, large-language models and foundation models) and to enable groundbreaking research in AI, computational infrastructure is essential. A university-wide investment in high-performance computing resources, including hardware accelerators and cloud-computing platforms, would provide researchers and students with the computational power necessary to effectively tackle complex AI problems.
v. **U:** As AI continues to evolve rapidly and the demand for computational resources grows, ongoing and sustained financial support and creation of fair-use policies across units will be necessary to ensure that Yale remains at the cutting edge of AI research and education. We urge the university to subsidize costs associated with large-scale computing to make it affordable, accessible, and appropriately secure (neither too strict nor too open).

2. We must consider how to use AI as a tool (for AI-augmented research and practice) as well as a basic scientific enterprise in which YSPH develops ethical and accurate AIs. As public health scientists, our primary role is to develop novel AI tools that are both more generalizable and more sensitive to local variation and to excluded populations.

   A. In our experience, the current generation of public health PhD students, postdoctoral students, and junior faculty see AI as core to their work. It is essential for us not just to support grant writing, but also to mentor faculty in developing and applying their tools in the real world.
      i. **YSPH:** To stay competitive with our peers, we are initiating multiple new AI faculty lines at junior and senior levels in expected and unexpected departments. We will continue both to recruit these new faculty and support existing faculty in public health data science, with a focus on those who work across relevant disciplines (e.g., with experience in computer science, statistics, demography).
      ii. **YSPH + U:** Introducing faculty to each other—within our school and across schools and units—is essential to unlocking the power of AI innovation. We will continue convening and holding research-in-progress sessions, and welcome university-level engagement with the same.
      iii. We encourage consideration of cluster hires in health AI with a streamlined application process for units who wish to participate.

B. To use AI as a tool requires investment in workforce training—both for existing faculty and staff and for the public health leaders of the future. A diverse pipeline of public health–data scientists is essential to supporting equitable innovation and implementation of AI tools.
    i. **YSPH:** We are investing in acknowledging and incorporating AI across our core and elective courses as well as enhancing our delivery of AI-relevant content to practitioners in the field (through short courses, Office of Public Health Practice–sponsored workforce training, etc.).
    ii. **YSPH:** With the arrival of our new senior associate dean, we will continue her NIH-funded Big Data Summer Institute dedicated to training the next generation of diverse data scientists, including those in the Connecticut community. We look forward to renewing this application in the future.
    iii. **YSPH:** We are currently fundraising for a dean's research fund to incentivize cross-disciplinary work on emerging health issues. AI-related research will be a major focus of this effort.
    iv. **U:** The Poorvu Center for Teaching and Learning will continue to be a primary partner and resource for us; we encourage and support continued partnership with it on best practices for AI in education.
    v. **U:** As we consider how to accelerate cross-sectoral collaboration on this emerging, culture-changing tool, we encourage a consideration of new models of faculty funding to permit easier collaboration (recognizing that our YSPH faculty are among the minority at Yale in relying on "soft money" for 70 percent of their twelve-month salary, and also recognizing the tremendous impact of directed pilot funds).

C. Public health faculty have always been leaders in the development of new methods for studies of health risk, health promotion, and health outcomes. AI is becoming an essential tool for this work, and we look forward to continuing to serve as international leaders in the use of novel AI methods (ranging from enhanced recruitment, to improved data collection and analysis, to improved communication with cohort or trial participants).
    i. **YSPH:** We will engage in discussions with our Yale Center for Analytic Science about novel methodologies to enhance the efficiency of design and performance of trial analysis. This work should include the Yale Center for Clinical Investigation as a partner.
    ii. **U:** In addition to our school-specific convenings, we encourage university-wide convenings to share how students and faculty are using AI to push the boundaries of health science across economics, medicine, management, computer science, statistics, philosophy, and more as well as public health.

3. Finally, we must consider the ethical and policy implications of this technology for health. We are committed to growing our relationships with the private and public sectors to inform the translation of our technical expertise into effective society-wide change.

A. We have multiple existing partnerships and policy relationships with industry leaders, policy makers, and community organizations concerned about the ways in which AI will change the health landscape and regulatory affairs.
    i. **YSPH:** Under the leadership of our new senior associate dean and our new director of InnovateHealth Yale, we are exploring novel partnerships to enhance access to compute resources and to provide a glidepath toward implementation.
    ii. **YSPH:** We continue to grow our ability to translate science into policy through relationships with policy makers at the state, national, and global level as well as with nongovernmental organizations (such as RAND), which help drive policy in more organized ways.
    iii. **U:** The support of Yale Ventures, University Corporate & Foundation Relations, and the Yale Office of Sponsored Projects in creating and accelerating these relationships is critical not just for our school but for the entire university.
B. We remain concerned about the huge negative implications for AI and health communication.
    i. **YSPH:** We are pursuing collaborations not just with policy makers, but also with big tech to build the scientific corpus and then inform best practice on the issue of AI-accelerated health disinformation.
    ii. **U:** The support of Yale Ventures, University Corporate & Foundation Relations, and the Yale Office of Sponsored Projects in creating and accelerating these relationships is needed.

Through these internal and external areas of work, we see Yale School of Public Health working with others across the university to lead the global future of AI for health.

The Architecture of the Mind: Artificial Intelligence and Education in Yale College

Pericles Lewis

MARCH 18, 2024

The *Yale Reports* of 1828, influential documents in the history of liberal education, emphasize teaching students "*how* to learn."[1] In what became a famous passage, Yale President Jeremiah Day argued:

> The two great points to be gained in intellectual culture, are the *discipline* and the *furniture* of the mind; expanding its powers, and storing it with knowledge. The former of these is, perhaps, the more important of the two. . . . Those branches of study should be prescribed, and those modes of instruction adopted, which are best calculated to teach the art of fixing the attention, directing the train of thought, analyzing a subject proposed for investigation; following, with accurate discrimination, the course of argument; balancing nicely the evidence presented to the judgment; awakening, elevating, and controlling the imagination; arranging, with skill, the treasures which memory gathers; rousing and guiding the powers of genius. (7) [emphasis in original]

For the past two centuries, the faculty of Yale College—along with most theorists of liberal education—have followed the *Yale Reports* in emphasizing the priority of "discipline" over "furniture"—that is, we focus on teaching students how to think rather than filling their minds with information. This brief preface is intended to place the report of the Committee on Majors regarding Generative AI in Undergraduate Education in historical context.

The history of technological innovation in the field of knowledge has tended to support this preference for discipline over furniture. Ever since Socrates complained that the invention of writing had undermined the skills of memory, educators have bemoaned the loss of craft skills in the face of new technology. (Of course, we have Socrates's argument only because his student Plato wrote it down, in the *Phaedrus*).[2] Others have welcomed such innovations as progress. In the past century, the pocket calculator and the computer made the slide rule obsolete. Word processing programs diminished the importance of traditional skills, such as handwriting and spelling. The internet and the mobile phone seemed to deal a death blow to the "furniture of the mind" as it was

1 Faculty Committee, "Report of the Faculty, Part I: Containing a Summary View of the Plan of Education in the College," in *Reports of the Course of Instruction in Yale College* [*Yale Reports*] (New Haven: Committee of the Corporation and the Academical Faculty [Yale College]; printed by Hezekiah Howe, 1828), 14. https://www.yale.edu/sites/default/files/files/1828_curriculum.pdf. For the matters discussed in this preface, I am grateful both to the members of the Committee on Majors and to Associate Dean Andrew Forsyth.

2 See also Jacques Derrida, "Plato's Pharmacy," excerpted in *The Derrida Reader: Between the Blinds*, ed. Peggy Kamuf (New York: Columbia University Press, 1991), 112–42.

argued that students could look up any facts they liked anytime and would have the resources of many libraries at their fingertips.

Throughout this period of change, the distinction between the discipline and the furniture of the mind has in fact remained relevant, but the borders between discipline and furniture have proven themselves ever more blurry. Is knowledge of biochemistry primarily about memorizing the names and relationships of amino acids, peptides, and proteins? Or is it about understanding the chemical structure of life and perhaps being able to intervene in it? We continue to teach students how to think critically, but we also need to teach them enough basic knowledge that they have something to think critically about. The introduction of large-language models (LLMs) and other new forms of artificial intelligence draws particular attention to the fact that knowledge is not just a string of facts (or words) but is a structure. Apparently, artificial intelligence (AI) tools such as AlphaFold 2 can model new protein structures with great precision and turn what might once have been the work of an entire PhD thesis into a one-hour project. What does the successful biochemistry student need to know to use these tools? She needs an understanding of the structure of biochemical knowledge—one might say the "architecture" of that knowledge, which is perhaps where furniture meets discipline. This architecture is a rudimentary form of what the Yale Task Force on AI has called "domain expertise."

In response to requests from the trustees and the provost, the Yale College Committee on Majors this year asked all departments to discuss the role of AI in undergraduate education. It seems likely that new forms of AI will have at least as dramatic an impact on the creation and dissemination of knowledge as did the worldwide web in the late twentieth century. We do not yet know how it will transform undergraduate education. As with the introduction of many new technologies, there is a considerable amount of fear, or at least discomfort, among those who underwent lengthy training in other technologies. As with those earlier technologies, it is likely that the key skills we teach in a liberal education will remain essential: analysis, argument, judgment, weighing of evidence, imagination. But our methods of teaching students how to learn will have to evolve.

Current State of Faculty Engagement with AI

Faculty have spent a good part of the last eighteen months thinking about assessment in the world of AI. Take-home assignments seem less relevant when AI can solve problem sets or write passable essays. Faculty have sometimes switched to in-class assignments, but this takes away valuable class time. Oral examinations would be a good way to judge how much a student has learned, but they are time consuming. Furthermore, students will need to know how to use AI tools in their working lives. The college has encouraged each faculty member to provide guidelines on the use of AI in coursework, and the Poorvu Center for Teaching and Learning has provided useful models of such guidelines.

Some Yale faculty are experts on AI; others seek guidance to a degree that is perhaps unprecedented. The Poorvu Center has provided useful guidance and become the go-to place on campus for information about AI. This is particularly helpful for faculty for whom AI is not a research subject.

Just as faculty involved with the AI Task Force have indicated an interest in centralized training for how to use AI in their research, we should develop a set of modules to help faculty understand how AI might affect undergraduate education.

While avoiding boosterism, we should recognize that AI presents some useful opportunities for undergraduate education and student affairs. We are already making use of AI in advising students on career strategy. Some faculty have begun to develop AI tutors for their courses—and some students have programmed their own AI tutors. It would be worthwhile to develop the skills of the Poorvu Center staff or other centralized resources to help faculty who want to learn how to undertake such experiments on their own.

Most crucially, the curriculum should develop rapidly to recognize that AI will be a part of every student's "toolbox," indeed a part of what every educated person will need to know. The undergraduate curriculum should have entry points for students coming to Yale with a range of backgrounds in the use of AI.

The following report from the Committee on Majors shows the unsettled nature of current faculty discussion about this topic. From the current fermentation may useful new proteins emerge for the future of undergraduate education.

Committee on Majors

March 7, 2024

Generative Artificial Intelligence in Undergraduate Education: Faculty Conversations, 2023–2024

In November 2023, Yale College's Committee on Majors invited chairs and directors of undergraduate studies to hold faculty conversations about current and future uses of generative artificial intelligence (AI) in undergraduate education. As of March 7, 2024, twenty-five majors—broadly representative of the disciplines taught—have shared reports of faculty discussions. Appendix 1 lists the majors.

The committee asked majors to discuss how faculty learn about AI; how AI is changing or will change course content, assessment, and pedagogy; and if AI tools or methods are being taught. Appendix 2 lists the questions asked.

Learning about AI

Faculty were first asked how they learn about AI. In addition to official guidance received from the provost, most majors report that faculty learn about AI from the Poorvu Center's website and workshops, and also from visits to faculty meetings or class sessions in introductory courses by the center's staff members.

Other sources of information for faculty include nonspecialist publications (e.g., *Chronicle of Higher Education, New York Times*), faculty members' own experimentation, and conversations with colleagues. More than one major also mentioned sessions conducted by the Center for Collaborative Arts and Media and the Center for Language Studies.

Some majors reminded the committee that their faculty members research and teach AI. Others stressed faculty members' lack of expertise.

Course Content

Faculty were asked if and how AI is changing their majors' course content.

Some majors have courses on AI. The content differs. Computer science (CPSC) engages "technical" questions—e.g., CPSC 488 AI Foundation Models—but also associated topics, such as AI and intellectual property rights, accountability, accessibility, privacy issues, and the effects on the environment. CPSC courses are being reorganized to give a better mathematical foundation to student understanding of AI and avoid redundancy among courses. The drolly titled CPSC 170, Artificial Intelligence for Future Presidents, is designed for nonmajors.

Some arts, humanities, and social science majors have courses or sections engaging AI topics (e.g., ARCH 332 Cultural AI: Machine Vision, Art, and Design; ENGL 114 Writing Essays with AI; GLBL 5065 Intro to AI: From Turing to ChatGPT; PLSC 338 AI and Democracy).

Some courses are adding substantive content on AI to courses. MB&B's 435a/635a, for example, Quantitative Approaches in Biophysics and Biochemistry, has four new lectures on data science and machine learning. Others discuss ethical uses of AI, including in scientific publications.

Other majors report changes to course content by instructional use of AI. East Asian Languages and Literatures (EALL), English (ENGL), French (FREN), and Statistics and Data Science (S&DS) report class discussions of AI output, often to show AI's limitations or the differences between AI and human products. Molecular Biophysics and Biochemistry (MB&B) suggests, however, that providing a meaningful critique of an AI-generated product—as happens in some humanities classrooms—is often too sophisticated a task for undergraduate learners.

A Theater and Performance Studies (THST) class considers style and creativity through listening to an AI-generated track of a song ostensibly "by Ella Fitzgerald." It's a song she never sang. Anthropology (ANTH) notes the importance of identifying "fakes."

More than one major reported an interest in bias and AI (e.g., African American Studies [AFAM] and Education Studies [EDST]), including AI's exclusion or under-representation of non-English sources. Although not directly stated by many, several reports implicitly assume that students should be taught how AI operates, in part to understand its potential limitations and biases.

Several majors reported that neither do they engage AI nor do they immediately intend to.

Assessment

Faculty were asked if and how AI is changing their assessment of student learning.

Many majors report that they do not use AI to assess student learning.

Some majors report that they explicitly prohibit students from using AI in assessed work. Others note the Poorvu Center's advice that syllabuses indicate how AI can and cannot be used. Implicit in several reports are worries about students using AI without permission and, in some fields, the difficulties of identifying plagiarism. Chemistry (CHEM) reports that it knows that AI does "too well" at answering its current problem sets in introductory courses and Physics (PHYS) that AI can already sketch out answers to its take-home examinations and will presumably get better and better at doing so.

Many majors report changes to assessment practices, with a new emphasis on in-class assessment, including in-class written tests and oral presentations (e.g., Film and Media Studies [FILM], FREN, German [GMAN], EALL, Mechanical Engineering and Materials Science [MEMS]). Some S&DS courses now allow AI to be used in examinations, with corresponding changes to the questions asked. Other majors note changes to writing assignments intended to make use of AI less effective and efficient—for example, through requiring multiple stages in the development of a paper or students' explanations of their choices. Some forms of assessment have been retired; an introductory ENGL class no longer asks students to produce verse in Popean couplets. AI can do that, if not wholly successfully:

In classrooms' realm, where knowledge takes its stand,
Assessment wields the rod, a guide's firm hand.

(ChatGPT 3.5, response to "Make a couplet in the style of Alexander Pope that explains the purpose of student assessment," March 7, 2024)

Not mentioned in majors' reports, but of possible concern, is whether increasing use of course time for in-class assessment will negatively reduce the amount of class contact time used to teach content. Similarly, take-home tests have hitherto usefully directed students' time outside the classroom.

Pedagogy

Faculty were asked for examples of how they anticipate that AI will provide new pedagogical opportunities. Majors report significantly different assessments of AI's possible contribution to pedagogy. Economics (ECON), for instance, reports that the "opportunities offered by generative AI in the teaching of economics far outweigh its downsides." An introductory EDST course teaches students how to use AI well: use it to conduct preliminary research, collect information, edit, tutor yourself on material, and provide summaries. Architecutre (ARCH) suggests that AI is already affecting the way students are working, so teaching must adjust accordingly, whether to embrace the potential of AI or resist its abuses. ENGL notes that AI is not yet producing the good clear writing taught by the major, but if it does, ENGL's pedagogy must change. Several majors report that they are not interested in using AI at present and are not optimistic about its uses: FILM, for instance, notes AI's potential to erode the creativity and precision of students' thinking processes.

Several pedagogical innovations are more anticipated than currently employed. These include gamification in the classroom; AI responsiveness to students' individualized learning, e.g., by providing tailored vocabulary lists in language courses; AI providing "tutoring" (PHYS); AI operating as an "auxiliary textbook," answering students' specific questions (MEMS); or AI generating practical problems to aid student learning (S&DS).

There are some worries about differing student preparation and access to AI. Some incoming students will have been taught how to use AI ethically and efficiently. Some will have come from schools that ban its use outright. Some students will have access only to free AI tools; others will pay for more powerful tools.

Teaching AI Tools and Methods

Faculty were asked if they anticipate teaching students to use new AI-based tools or methods.

Many majors anticipate doing so. Some majors mention specific tools, e.g., in MB&B, AlphaFold 2 predicts the structure of protein molecules with high accuracy—historically a barrier to research progress—and its use will likely be incorporated in several courses. FREN mentions DeepL, which provides reliable translations between modern languages.

Several majors—including Computer Science (CPSC), ECON, PHYS, S&DS—mention that a significant use of AI is to streamline programming and debug code, with implications not least for data analysis, mathematical modeling, and scientific word processing. Opinions differ on the effects on student learning. ECON sees no intrinsic benefit in students learning a particular programming language. S&DS notes that if coding is faster and easier, time will be freed up to teach other topics. CPSC strikes a note of caution: learning a program may help students to develop their algorithmic thinking and solve problems efficiently.

Others point to more general uses for AI, including for citation management (ECON, ENGL), to translate (THST), and to summarize primary and secondary sources (Ethnicity, Race, and Migration [ER&M]), again with disagreement about whether these will aid or hinder student learning. Biomedical Engineering (BENG) expects that AI will affect the generation of biomedical data in many settings. If and when it does so, AI should be taught in BENG courses.

Other Themes

Faculty were invited to share with the committee other themes that arose in their conversations. Some majors ask for greater guidance from the university and suggest that a systematic approach to educating the faculty is needed (ER&M, GMAN). Faculty levels of knowledge differ greatly (noted MEMS), with many knowing less than their students. In addition, in any guidance, the university should not focus solely on text; specific issues arise in fields in which images are important (e.g., ARCH, BENG).

Some cautioned about the approach being too instrumental. There is a danger of under theorizing AI. If teaching the building blocks of AI or its use is the first level, then second-level ethical analysis of using AI is important but insufficient. Third-level consideration of what AI "means" for human life and society is also needed.

Differing views are expressed on the overall benefits of AI and on whether the university should promote AI's use. CPSC suggests that research to improve and extend AI methods will continue at Yale and elsewhere, regardless of any opposition. ANTH worries about AI decreasing human creativity. ECON suggests promoting AI's use in student learning, while FREN cautions against this. ENGL notes possible faculty demoralization if faculty spend significant time grading AI-generated products.

Some majors imagine near-term curricular uses or implications. PHYS suggests that soon enough an AI-driven platform could help students learn mathematics or related skills in advance of taking college-level classes. EALL suggests that, if translation tools continue to improve, then foreign language requirements may change where languages are not essential to the scholarly task at hand. Will an East Asian–language graduate student need to learn modern European languages if secondary scholarly output in various languages can be sufficiently well translated by AI?

More than one major noted and welcomed general faculty discussion about what should be taught and how, irrespective of the specific prompt to consider AI in undergraduate education.

Appendix 1. Reports of Discussions Received

Reports were received from the following majors. Some departments support more than one major and shared a single report.

Humanities and the Arts

- Architecture (ARCH)
- Art (ART)
- East Asian Languages and Literature (EALL)
- English (ENGL)
- Film and Media Studies (FILM)
- French (FREN)
- German (GMAN)
- Theater, Dance, and Performance Studies (THST)

Social Sciences

- Anthropology (ANTH)
- Economics (ECON)
- Psychology (PSYC)

Biological Sciences

- Molecular Biophysics and Biochemistry (MB&B)
- Molecular, Cellular, and Developmental Biology (MCDB)

Physical Sciences

- Chemistry (CHEM)
- Physics (PHYS)
- Statistics and Data Science (S&DS)

Engineering

- Biomedical Engineering (BENG)
- Computer Science (CPSC)
- Mechanical Engineering (MENG)

Interdisciplinary

- African American Studies (AFAM)
- Computer Science (CPSC) and Economics (ECON)
- Ethnicity, Race, and Migration (ER&M)

- South Asian Studies (SAST)
- Women's, Gender, and Sexuality Studies (WGSS)

Certificate Program

- Education Studies (EDST)

Appendix 2. Questions shared

Chairs and directors of undergraduate studies were asked to use the following questions to generate faculty conversation.

- What resources on campus or other professional resources have your faculty utilized to learn about generative AI?
- Please describe if generative AI has changed the way your department:
  - Provides content in your courses. For example, do you provide AI examples in your courses, including ways to show limitations and biases of AI?
  - Assesses student learning in your courses.
- Provide examples of how your department anticipates that generative AI will provide opportunities for new pedagogy. For example, do you anticipate AI contributing to personalizing student learning, gamification (adding gamelike aspects to your course materials), providing feedback to students, tracking student progress, and/or adjusting material difficulty? If so, please describe.
- If generative AI tools impact research in your discipline, do you anticipate changes to the curriculum of your major to teach students to use new AI-based tools or methods that utilize AI?
- Are there other aspects of how generative AI might impact your major that you would like to communicate to the Committee on Majors?

Collections and Scholarly Communication Division Report
Yale Task Force on AI

APRIL 3, 2024

The rapid rise and adoption of generative artificial intelligence (AI) over the past eighteen months has proven to be a disruptive force in many areas of society, including the focal points of Yale University's mission. Market analysis by Gartner puts generative AI as the technology which will most impact and improve productivity in the coming year, especially when backed by large-scale data managed as a knowledge graph — a database of relationships among described entities, rather than a set of documents or tables.[1] In 2023, through the Cultural Heritage/Information Technology (CHIT) collaboration, Yale launched LUX, a knowledge graph that makes the collections of the Yale University Art Gallery, the Yale Peabody Museum, the Yale Center for British Art, and the Yale University Library discoverable in a single place for the first time, and that provides extensive contextual knowledge about the related people, places, events, and subjects. Backed by this knowledge graph, generative AI offers an unparalleled opportunity for us to activate centuries of work in collecting, preserving, and describing our more than 30 million items of cultural and natural heritage. By working with the collections at a scale beyond human capacity, we can provide a unique experience for our students, facilitate the creation and delivery of unique cross-disciplinary courses, and make previously hidden knowledge globally accessible to researchers.

An engineer at OpenAI wrote that, after training many different models with different settings and algorithms, all the systems with the same dataset eventually converged to the same functionality.[2] It is the data that provides the value for the AI, and all other factors create a somewhat efficient process to approximate the knowledge in that data. We can increase our leadership status within the cultural heritage–knowledge industry by ramping up our investment in the creation and integration of high-quality data with generative AI. That data includes structured metadata, 2- and 3-dimensional images of physical objects, scientific datasets, and beyond.

Generative AI can dramatically accelerate existing transactional tasks of staff and faculty, allowing time for more strategic and innovative work. By integrating this technology into core processes throughout the Collections and Scholarly Communication (C&SC) division, we will be able to use our limited human resources more efficiently to perform tasks that only such professionals can accomplish. For example, the basic cataloging of a new book acquired by the library could easily be performed by an AI system with access to images of the title page, copyright statement, back cover or abstract, and table of contents. Digital-accessibility requirements for images of artifacts at the art gallery can be met with the same approach. Instead of having people create these records manually,

1 Lori Perry, "30 Emerging Technologies That Will Guide Your Business Decisions." Gartner (February 12, 2024). https://www.gartner.com/en/articles/30-emerging-technologies-that-will-guide-your-business-decisions

2 jbetker, "The 'it' in AI models is the dataset." Non_Interactive — Software & ML (June 10, 2023). https://nonint.com/2023/06/10/the-it-in-ai-models-is-the-dataset/

they can instead review the output of an AI-based system. Generative AI will dramatically accelerate the description and discoverability of our collections.

Use of generative AI will transform access to the collections, ensuring that our extraordinary museums and library collections are readily accessible for purposes of cross-disciplinary teaching, learning, and research and that they are responsibly stewarded now and into the future. We envision a near future where diverse audiences, both internal and external, can engage digitally with the collections through AI-mediated tools and knowledge bases. These tools will universally improve access to the cultural heritage knowledge and items that are curated, preserved, and made accessible at Yale. This will take many forms, including natural language–based search interfaces, automatic summarization of full-text search results, and the discovery of valuable audio-visual content, research datasets, or other content that has previously been impossible to engage with beyond basic descriptive metadata.

We have pioneered basic supervised machine learning (or nongenerative AI) techniques across the division, including to predict material measurements of photographic paper in the Lens Media Lab in the Institute for the Preservation of Cultural Heritage (IPCH) to segment and categorize images of natural history specimens in the Peabody, and to disambiguate place names for reconciling across collection datasets within LUX. We have engaged with external organizations around their plans in this realm, including peer universities, national museums and libraries, and far beyond. Approximately sixty organizations have approached us about adopting LUX for their collections, and generative AI will enable broader dissemination of our resources as well as discoverability of cultural heritage collections worldwide.

We intend to use AI products to meet mission-critical goals in the division starting with a pilot to transcribe and translate digitized handwritten or printed notes and documents about the ownership history of our collections in all languages. From the text extracted from the images, we can further recognize entities such as people, places, objects, and events and then summarize the information to enable researchers to be highly efficient in piecing together the complex patterns of acquisitions, gifts, loans and, perhaps, even thefts of the artifacts that we look after. Understanding the history of our collections is a legal and moral obligation that AI will dramatically accelerate. This information will be made more accurate through integration with LUX and the results fed back into it in a virtuous cycle of knowledge acquisition.

Data, especially but not exclusively within the arts and humanities, often suffers from bias, harmful language, inaccuracies, imprecision, and incompleteness. This makes it challenging to have a machine without human experience engage directly with collections data. Examples include when the Smithsonian Institution asked an AI to describe an image of slave shackles, it responded with "bracelets" in the domain of jewelry; less harmful, but no less embarrassing, the Swedish National Museum responded with a query for "baseball" with images of jugs, as the AI had interpreted baseball as including the notion of "pitcher," and was unable to distinguish between the homonyms. Even Google had to pull back the image-generation capabilities of Gemini after it generated highly

inappropriate images of historical figures.[3] We have the data, expertise, and professional networks to help move the field forward in improving semantic accuracy, reducing inadvertent harmful language, and reducing the "hallucinations" from which AIs suffer by researching innovative and reproducible approaches to detecting and preventing these reputation-damaging responses.

The Collections and Scholarly Communication division strongly supports universal and equitable access to the appropriate suite of generative AI tools for students, staff, and faculty. In conjunction with the Poorvu Center for Teaching and Learning, we will provide leadership and education for the Yale community to promote understanding of when generative AI is appropriate, how to use it, which tool is the most effective for a particular task, and to provide a supportive and considerate environment in which experiments and learning can take place. A shared, cross-disciplinary service laboratory staffed with appropriate expertise and with sufficient resources to get people started would further this vision, following the model in place for the Digital Humanities Lab in the Yale University Library.

By leveraging our existing strengths — our collections, our knowledge graph, and our staff's expertise and professional networks — we will continue to provide leadership and guidance for the cultural heritage domain as a whole. The challenges outlined above are faced by all collecting organizations, and providing a responsible, culturally sensitive, and well-integrated solution will improve the world today and for future generations.

Appendix A: Institutional Policies, Activities, and Resources

The Collections and Scholarly Communication division acknowledges several gaps uncovered by the rapidly changing AI landscape that should be addressed with new policies and resources at the institutional level.

Policies

1. The digital divide is widened dramatically by the time-saving capabilities of these tools, allowing those with access to and understanding of generative AI to learn, teach, research, and work significantly more effectively and efficiently. The availability of generative AI products should be treated in the same way as access to the internet, core information technology products, and the knowledge stored and licensed by the Yale University Library. We propose to help with improving the understanding of generative AI with a cross-disciplinary laboratory situated in the library, and in conjunction with the Poorvu Center.

2. The use of generative AI tools has been sporadic and unstructured. Some staff, who feel comfortable with technology, have jumped in with both feet. Others have not known whether they were allowed to, how much it might cost, or where to begin. A consistent and coherent policy around the appropriate use of these tools, with the resources and training available as per the above, would alleviate the concerns of many.

3 Adi Robertson, "Google apologizes for 'missing the mark' after Gemini generated racially diverse Nazis." *The Verge* (February 21, 2024). https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical

3. As generative AI technology continues to change and improve at whirlwind speed, all divisions and schools will need the agility and flexibility with budgets and staff to pivot and follow suit or seek alternative avenues for their domain. All schools and divisions will need access to resources with which to experiment and conduct pilot studies, without having to wait for the annual budget process to obtain those resources.

Resources

1. Predicting how many graphics processing–unit (GPU) or tensor processing–unit (TPU) cycles are needed is impossible given the flux in algorithms and data structures multiplied by the different products, paradigms, and usage patterns for batch processing of knowledge versus end-user engagement. Further, GPU and TPU hardware is constantly improving. Much like the central expertise and resources managed by the Yale Center for Research Computing, a central pool of processing cycles or available on-premises processors would accelerate innovation and allow the creation of new services or products supplemented by licensing processors in the cloud.
2. Additional resources should be made available for the digitization of physical materials for the express purpose of creating training datasets to improve the accuracy and breadth of generative AI solutions at Yale. While C&SC would benefit from the digitization of archival content held in the library and museums, all areas of the university have significant paper trails that likely would also provide rich sources of knowledge about their organizational mission and practices.

Artificial Intelligence in Clinical Practice
Report to the Yale Task Force on AI

Margaret M. McGovern, MD, PhD
Jason Fish, MD
Lee Schwamm, MD
Lisa Stump

APRIL 1, 2024

Artificial intelligence (AI) is already driving a revolution in healthcare and is poised for mass adoption, with applications ranging from robotic process automation to machine learning and generative AI. Algorithms for image interpretation and clinical decision support are already integrated into electronic health records and radiology interpretation systems and are entering the sphere of clinical documentation at a rapid pace. Applications to the Food and Drug Administration (FDA) for approval of software as a medical device (SAMD) are sky rocketing, and it is only a matter of time before nearly all healthcare workers engage with AI solutions on a daily basis. Many of these healthcare workers are and will continue to be unaware that AI is at work "under the hood" in the course of their daily transactions. At the same time, many healthcare workers express concerns about the use of AI and its potential to dehumanize care, "hallucinate" false results, and put their jobs at risk.

Based on the prevalence and growth of AI solutions in clinical practice, we have an obligation to advance the safe, equitable, and effective adoption and deployment of these solutions. In routine business operations, AI can enhance human productivity and automate many mundane tasks traditionally requiring human effort. This shift will obviate the need for humans to perform these more mundane tasks, which will enable the workforce to focus on performing higher-skilled tasks. In clinical decision making, AI should be used to augment and support human clinicians, not replace them, and clinicians should be trained to evaluate critically the output of AI solutions.

However, few frameworks and little guidance exist to support the responsible implementation of AI in clinical practice. Several national collaboratives of industry and academic partners are taking shape to begin to grapple with these complex issues. Model transparency, diverse training data, and post-implementation evaluation are emerging as central themes. Yale has an opportunity to become an authoritative voice in all of these domains, leveraging expertise across the campus in educational design, ethics and social impact, and scientific methods in implementation and evaluation.

With our partners in the Yale New Haven Health System (YNHHS) and Yale Health, where the integration of these tools into the digital ecosystem of the electronic health record, imaging systems, laboratory systems, and the diverse clinical training and care environments occur, Yale can amplify the impact of its contributions in AI adoption on the millions of patients cared for every year. This will complement the important work being proposed by the task force members focusing on clinical

research, basic research, and public health, including the computational infrastructure needed to deliver on the promise of AI.

In the pursuit of these opportunities, Yale University and YNHHS are positioned for leadership and bolstered by several strengths and assets:

1. A unified, systemwide, longitudinal electronic health record (EHR), which encompasses over 4 million unique patient records over decades of care encounters.

2. In addition to the EHR, our data assets combine imaging, lab, pathology, bedside alarms, financial, and operational data. The data architecture is based on standard data models and nomenclature systems (e.g., RxNorm, LOINC, and OMOP). Together, this provides a standardized data model framework with rich and diverse information for model validation, outcomes analysis, and performance-improvement efforts both within Yale and in federated queries, together with other organizations. We recently enabled the Cosmos module in EPIC, which supports federated deidentified queries of over 200 million patients across all Cosmos-participating EPIC sites.

3. The patients served by YNHHS, Yale Health, and Yale School of Medicine (YSM) are diverse and reflective of the ethnic and demographic makeup of the U.S. population as a whole. This facilitates the generalizability of models and allows for the development and assessment of clinical AI models for effectiveness and potential bias across a wide range of patients.

4. Our clinical faculty provide care in a variety of delivery models and settings (e.g., complex inpatient care, comprehensive ambulatory care, home care, emergency medicine, urgent care, dental care, remote monitoring, and virtual care) and payment models (e.g, self-insured, commercial, governmental, and self-pay). This enables living laboratories that span a broad continuum, and it also allows for the development and testing of clinical AI models.

5. Broad expertise exists among faculty in YSM, YNHHS, and across the university in the ethical aspects of introducing technology into clinical care, which includes the fair and ethical design of AI and large-language models (LLMs), community engagement and participation in our frameworks for ethical care delivery in our patient and family advisory councils at YNHHS, and legal, ethical, economic, regulatory, and policy experts across the university, health system and community groups.

6. The Yale Center for Healthcare Simulation provides an ideal opportunity to introduce AI and technology-enabled care and training into a living laboratory, where it can be vetted prior to implementation in real clinical encounters and where clinicians can practice and develop competency with these new tools.

To capitalize on these opportunities and position Yale as a national and international leader, we propose the following areas of focus and investment:

1. Robust and well-designed IT infrastructure will be needed to support these efforts across the university to provide the needed capacity for data storage and computational analysis. Bringing together the clinical, financial, and administrative data in a reliable, secure, and accessible ecosystem will require joint efforts between the university and YNHHS. Foundational to the success of this work is a federated and comprehensive approach to user identity and access management, managed collaboratively via the cloud-based Microsoft Azure active directory across the two organizations. This work has begun and requires additional focus and resources to mitigate current challenges and enable breakthrough innovations. Key components of this infrastructure include:

- In collaboration with YSM, YNHHS and the Yale Center for Research Computing (YCRC), establish a robust HIPAA-compliant infrastructure to support the AI lifecycle, ensuring a secure compute environment that is accessible to university-wide users, has strong access and identity management controls, and can run against real-time clinical data.
- Scalable, high-performance servers or specialized hardware — graphics processing units (GPUs) — to handle the computational demands of AI and with redundancy to protect against data loss or interruptions in hardware performance, as described in detail in the YSM AI report.
- Sufficient and scalable storage capacity to manage structured data (e.g., physiologic and laboratory values, genetic data, clinical attributes, demographic, and financial) and unstructured data (e.g., free-text clinical documentation, radiology, cardiology, and pathology images, established and novel sensor output).

2. Education and training with AI tools and new care-delivery models is an area where Yale is poised to lead the nation. This includes:

- Redefining the nature of medical education in a world where every teacher, learner, and trainee engage with AI on a daily basis. This work would be conducted in coordination with Jaideep Talwalkar MD (assistant dean for education, medical education) and Jessica Illuzzi, MD (deputy dean for education, Harold W. Jockers professor of medical education, and professor of obstetrics, gynecology, and reproductive sciences).
- Establishing the scope, format, competency, and evaluation methods unique to clinical practice and ensuring that these tools augment rather than diminish human performance.
- Creating and delivering educational content that can power the next generation of clinical and preclinical training in collaboration with leaders in graduate medical education, nursing and allied health professions, and the Center for Medical Education.
- Build and disseminate training materials and programs to support the workforce in adapting to a workplace that incorporates AI alongside humans and will be able to engage with AI from a position of knowledge and awareness of risks.

3. Development of comprehensive, personalized educational materials for patients to help them navigate the healthcare system, express their expectations for care, and enable them to overcome barriers to adherence with care plans based on social drivers of health and other factors.

In addition, we need to educate and reassure patients about the use of their health data in the training of AI models, the safe and responsible use of AI in their health journey, and how to be more knowledgeable consumers of healthcare in understanding the limits and value of AI.

4. A comprehensive evaluation framework is needed to enable the responsible, high-quality implementation of clinical AI in practice and to maximize usability and deliver exceptional patient and clinician experiences.

- The ethical dimensions of AI in clinical practice are numerous and ill-defined. Expertise across the university is needed to ensure that the core ethical dimensions of fairness, transparency, utility, and equity are formally and explicitly incorporated into the evaluation framework of any newly proposed and implemented AI solutions.
- Bringing together expertise in implementation science, the social sciences, and mixed methods approaches will enable Yale to establish thoughtful and pragmatic approaches to this complex task to inform the national conversation, as is also described in the YSM AI report.
- Yale University has an opportunity to amplify its voice in partnership with digital leaders in YNHHS, who serve as members in several national collaboratives devoted to the assessment and responsible implementation of clinical AI (e.g., www.coalitionforhealthai.org, www.validai.health, Trustworthy & Responsible AI Network—TRAIN).

5. Creation of living laboratories (in both simulated and actual clinical settings) in the learning healthcare-system environment shared by Yale University and YNHHS will allow for the controlled and phased implementation of AI into clinical practice with a virtuous and iterative cycle of learning prior to system-wide deployment. Yale Health, with its clinical expertise, population-health responsibility, and payer-provider status provides a unique and ideal location for these living laboratories. Other opportunities exist within the health system in partnership with clinical faculty in YSM in areas such as the Yale Center for Healthcare Simulation in emergency medicine as well as in pediatrics, geriatrics, behavioral health, cancer, and other clinical centers.

6. Enhanced engagement in inside-out and outside-in innovation efforts to both develop and codevelop new commercial applications and clinical AI solutions can improve health outcomes and drive research. These activities are already underway in the collaboration between Yale Ventures, Yale faculty, and various innovation centers across the health system and university. Both the health system and the university are actively engaged with the Advanced Research Projects Agency for Health (ARPA-H) through grant applications and membership in the investor catalyst, respectively, and the university will be hosting a first ever, New Haven–based innovation incubator with Techstars. Yale can lead in this area by defining the clinical practice domains that are a priority for AI-powered innovation, coordinating the various centers of innovation to incentivize inventors to tackle these priority challenges, and leveraging the living laboratories as evaluation centers to inform product design and implementation.

Appendix 1

AI in Clinical Practice Panel members

- Cochair: Margaret McGovern, MD, PhD, CEO, Yale Medicine; Deputy Dean for Clinical Affairs, Yale School of Medicine; Executive Vice President & Chief Physician Executive, Yale New Haven Health System
- Cochair: Jason Fish, MD, CEO, Yale Health
- Lisa Stump, MS, Lecturer; Senior Vice President and Chief Information Officer, Yale New Haven Health System and Yale Medicine
- Lee H. Schwamm, MD, Associate Dean, Digital Strategy & Transformation, Office of the Dean, Yale School of Medicine; Professor in Biomedical Informatics & Data Science and Professor of Neurology, Yale School of Medicine; Senior Vice President & Chief Digital Health Officer, Yale New Haven Health System
- Sanjay Aneja, MD, Assistant Professor of Therapeutic Radiology; Director of Clinical Informatics, and Director, Medical School Clerkship, Therapeutic Radiology; Assistant Cancer Center Director, Bioinformatics
- F. Perry Wilson, MD, MSCE, Associate Professor of Medicine (Nephrology) and Public Health (Chronic Disease Epidemiology); Director, Clinical and Translational Research Accelerator (CTRA); Course Director, Interpretation of the Medical Literature; Codirector, Human Genetics and Clinical Research Core
- Michael Hoge, PhD, Professor of Psychiatry; Director, Yale Behavioral Health; Director, Yale Group on Workforce Development
- Rohan Khera, MD, MS, Assistant Professor of Medicine (Cardiovascular Medicine) and of Biostatistics (Health Informatics); Clinical Director, Center for Health Informatics and Analytics, YNHH/Yale Center for Outcomes Research & Evaluation (CORE); Director, Cardiovascular Data Science Lab (CarDS)
- Andrew Taylor, MD, MHS, Associate Professor of Emergency Medicine and of Bioinformatics & Data Science; Director of Artificial Intelligence and Data Science, Emergency Medicine
- Harlan Krumholz, MD, SM, Harold H. Hines, Jr. Professor of Medicine (Cardiology) and Professor in the Institute for Social and Policy Studies, of Investigative Medicine, and of Public Health (Health Policy); Founder, Center for Outcomes Research and Evaluation (CORE)
- Madeline Wilson, MD, Chief Campus Health Officer and Chief Quality Officer, Yale Health
- Nanci Fortgang, RN, MPA, CMPE, Chief Clinical Operations Officer, Yale Health
- Peter Steere, RPh, MBA, Chief Operating Officer, Yale Health

- Hema Bakthavatchalam, CPHIMS, Director of Information Technology, Yale Health, and Director, Clinical IT Optimization at Health Sciences Information Services
- Wies Rafi, PhD, MSc, Associate CIO, Health Sciences IT, Yale University ITS
- Mary Hu, MBA, Associate Dean for Communications, Yale School of Medicine

AI in Clinical Practice Resource Recommendations

In addition to resource recommendations that are spelled out in the YSM AI report, we estimate that additional resources will be required to implement the integrations necessary within the YNHHS infrastructure and existing electronic health record and digital application systems. These include:

- Data and computer network architects and identity and access management engineers
- Prompt engineers, data scientists, language model specialists
- Dedicated application analysts and builders to integrate new AI models into real-time clinical applications
- Analysts to aggregate the data and support the build and ongoing monitoring of new AI model performance in the post-implementation phase
- Human-centered designers to create next-generation AI-enabled multimedia patient educational materials and delivery methods

AI in Operations

MARCH 27, 2024

The latest annual McKinsey Global Survey on the current state of artificial intelligence (AI) confirms the explosive growth of generative AI tools. Less than a year after the debut of many of these tools, one-third of leading firms and institutions are using generative AI in at least one business function and are evaluating how to expand. Yale Operations, too, is experimenting; now we need to accelerate progress.

Key Implications

1. **Enthusiastic support:** The Operations Leadership Team (OLT) is keenly interested in the potential of AI to transform administrative functions. Collectively, we are all intrigued with the benefits AI can bring to enhance decision making; optimize processes, foster a proactive problem solving; and improve customer service, including increased efficiency, accuracy, and the ability to make data-driven decisions.

2. **Disciplined investment:** We intend to leverage AI capabilities that are increasingly offered in new off-the-shelf tools like ChatGPT and GitHub Copilot or new enhancements in tools already used at Yale, such as Workday. The significant incremental university investments should likely focus on the academic and research opportunities.

3. **Leverage multi-industry best practices:** Looking ahead, we will invest the time to learn how other organizations, both in higher education and for-profit, are using AI. Many of our work activities are common, which should help identify use cases and accelerate the development of standard processes.

4. **Enhanced focus on data:** We acknowledge the existing challenges in the lack of common data definitions, uneven data governance, and silos that block institutional use. We endorse the current work underway to develop a university-wide data strategy and support organization.

5. **Build capability and broaden experimentation:** Many early adopters are already experimenting with AI across operations with specific pilots underway. We intend to sponsor training and develop programming to build skills and encourage use. Finding use cases across our four strategic initiatives — OneFinance, Recruiting, Learnings, and the Job Framework Redesign — is an immediate priority.

Campus-wide Opportunities

While our primary focus has been on the use of AI within our functions, the OLT acknowledges that we have a foundational role in supporting the use of AI across the university. Your input and guidance should provide us with the beginning of a roadmap, but we do anticipate that the next steps on a university-wide basis will likely include:

- Resolving infrastructure expectations (i.e. graphics processing unit hardware/services)
- Addressing access considerations
  - For different groups across campus?
  - Using which core suite of software and tools?
- Expanding training and exposure
- Clarifying guidelines associated with differing data classifications
- Advancing culture work surrounding institutional data

We anticipate working closely with the Yale Task Force on AI to align on the immediate and medium-term plan to support these campus-wide opportunities.

AI as an Accelerator in Operations

Looking ahead, we believe AI to be an accelerator to the realization of our overall goals of cultivating a culture of innovation and foresight, driving operational excellence, and enhancing the Yale experience for all stakeholders. Our aspiration is much bigger than just implementing AI technology. This journey is about more than just technology; it is about shaping a smarter, more responsive, and proactive Yale.

In this endeavor, we are committed to ethical AI practices, ensuring transparency, fairness, and respect for privacy at every step. We see AI as a partner in our mission, helping us to uphold the values and standards that Yale University proudly represents.

Opportunities

AI has the potential to significantly impact efficiency and the effectiveness of operations. Because most operational units are industry agnostic (finance, human resources, and facilities are universal), there is a bristling market responding to AI development across these functions. As a result, our emphasis will be on leveraging existing product and services integration of AI tools as our primary means of adoption. We are involved in numerous pilots in most of these areas:

1. **Personal productivity.** Individuals across operations are experimenting with the use of generative AI products in content generation. Initial memos, performance management, or drafting standard operating procedures and similar daily tasks are areas of potential use. In addition, productivity tools such as Microsoft Office and Teams, as well as integration into Zoom are providing reasonable first-draft meeting notes among other mundane tasks that ultimately enhance productivity. We continue to enable AI campus-wide capabilities as they meet security and privacy requirements, but also expect to consider operational user needs as we consider broader campus access requests.

2. **Call center/Support functions.** One of the areas in which early adopters have seen the most significant impact from the use of generative AI has been to support help desk and customer service inquiries. Using RAG (retrieval-augmented generation) methods against commercial large-language models are increasingly being used to create chatbots and copilot tools for both

improving the efficiency of help desk staff, but also in expanding hours of support through agentless chat. We have several pilots underway, notably AI-enabled chat through askyale.edu and copilot for call center staff for Yale Health.

3. **Software coding.** Several generative AI products have demonstrated the ability to write coding routines that accelerate some forms of custom code development. These have the potential to reduce workload and shift programming job roles over time. We are exposing the central software-development teams to these products and have begun to see some positive outcomes, with a pilot group of approximately twenty developers now actively experimenting with a variety of AI code–generation products.

4. **Image processing.** Image-processing tools are beginning to gain substantial capability. This has some practical implications to our work, notably parsing police-camera videos for specific investigatory needs, license plate recognition to improve parking enforcement, and an early prototype being used to examine refuse to aid hospitality in reducing food waste.

5. **Automate cybersecurity at machine speed.** AI is being used by bad actors to accelerate attacks, and the cybersecurity vendors are integrating AI into their defenses in a new arms race. These tools are necessary to combat hostile actions such as phishing and network intrusion. We are actively rolling out some products in this space to improve our automated response and reduce harms.

6. **Prediction and forecasting.** AI tools are embedded in many products that scan large amounts of data and return narrower results based on defined goals. This is most mature in areas such as employment-recruiting candidate screening, predictive maintenance models for facilities equipment, and financial forecasting. These requirements are being considered in upcoming projects, most notably in the operations talent acquisition (recruiting) project.

7. **Complex written materials analysis.** AI models are proving effective at summarizing complex contracts, policies, or similarly dense textual material and creating summaries or comparisons that can aid in efficiency of review or call out terms that may need additional review or revision. We continue to experiment with ingesting targeted data sets (i.e. financial policies) to see how AI may improve our ability to provide clearer guidelines and answer questions more efficiently.

8. **Process productivity.** A number of evolving AI tools can aid in the development of improved process from review of transactional data. These tools hold promise for simplifying process and procedural development.

9. **Data analysis.** Substantial opportunities exist to leverage business process automation and machine learning in combination with AI tools to better work with large-scale datasets in operations. These include areas such as space utilization, employee retention, and fraud detection. Pilots are actively underway in space utilization and fraud detection.

From experiences to date, we see three major challenges for the adoption of AI within Operations: First, training and exposure has been modest, and gaining confidence with these tools requires

a willingness to explore and learn in a low-risk environment. If we wish to accelerate adoption, teams will need some structured opportunities to engage. Second, improving our data practices will be crucial to progress with AI. Our quality-assurance practices are inconsistent across functions, and AI is reliant on clean data to derive consistent and predictable results. Our fragmentation and divisional segmentation of data also stands as a major barrier to harnessing the effective use of institutional data, and we will need to begin to evolve our practices to maximize the utility of AI investments. Third, inspiring a culture of experimentation and safe mistakes will challenge our historical practices. The AI market is extremely fast moving, with several years of vying solutions and emerging products ahead. We will inevitably make mistakes, and some investments will ultimately show poor returns. Organizations seeing initial success are experimenting and "failing fast," learning from those experiences, and trying again. Our institution's administrative 322-year history is highly deliberate. Success will require some departure from that culture and adopting a more research-like culture.

Recommended Investments

- Amplify investments in policy related to data strategy and governance, improving quality and consistency of institutional practices. We are at a stage with our work on analytics where increasing our focus on institutional data, investing in roles that focus on data quality and governance is a necessary precursor to our ability to materially advance the use of artificial intelligence tools within core operational functions.
- Update guidance to encourage experimentation and celebrate learning from inevitable failures. Our initial guidance may have created a more chilling environment than intended. Within Operations, most seem to have focused on the "do not" rather than the "do," and are awaiting clearance to engage with the use of AI. While that is beginning to be corrected through discussions and pilot activities, as we formalize solutions available to the campus, it will be important to shift the language to be more encouraging of use and experimentation.
- Design and build relevant training and exposure with a core cadre of users to drive learning and innovation. One of the most overarching comments across all working groups has been the desire for more exposure through hands-on workshops, sanctioned educational material, and recommended reading. Due to the overwhelming amount of press on the topic, curiosity is high, but the actual user group remains a relatively small set of early adopters and technologists. To advance into areas of high value will require a much broader immersion for a wide swath of staff.
- Selectively invest in skills development required to both leverage and reinforce AI models. We are building targeted skills on the margins today, and formal positions have yet to be established to support broad community needs in AI. It is likely this would best take the shape of acknowledging new skills within existing teams rather than designating an AI-specific work group.
- License access and support for relevant tools, prepare for product segmentation and cost growth as the market develops, transitioning toward operational standards over time. We are likely to see many tools respond to operational needs, not one. Generative AI tools are a part of this equation,

but to harness use across operations also involve application programming interface (API) integration and a series of tools for data analysis, many of which will best be served as extensions on existing major applications. As with all innovation, this early period will have broad diversity. As the market settles over the next several years, determining similar needs for access and support among schools and units versus truly unique needs will help to manage licensing proliferation and cost growth.

- Strengthen process centricity of Yale, improving consistency of practice will yield broader benefit and opportunity. Yale continues to grow in scale, and necessarily process maturity is evolving. While AI can help to expose commonalities in disparate processes, making use of these advances ultimately depends upon reaching broader consensus and determinate standards. As institutional scale approaches the $10 billion mark, this will be increasingly important for containing institutional administrative-cost growth.