Sen. Doc. No. 24-087

SPECIAL REPORT

OF THE

JOINT TASK FORCE ON GENERATIVE AI

FINAL REPORT

FINAL REPORT OF THE JOINT TASK FORCE ON GENERATIVE AI

FINAL REPORT

Presented at the

830th Regular Meeting of the Faculty Senate

May 2, 2024

Membership:

Fran Berman
Steve Bischof
Jeffrey Blachard
Allison Butler
Claudia Donald
Laura Haas
Farshid Hajir
Claire Hamilton
Kevin Kerwood
Marc Liberatore
Chris Misra
William Murphy
Jennifer Normanly
Anthony Paik
Laura Quilter
Chase Randall
Matt Rattigan
Erin Schadel
Michelle Trim co-chair)
Jackie Urla
Ruth-Ellen Verock-O'Laughlin
Loren Walker
Tilman Wolf (co-chairs)
Ethan Zuckerman

I. Charge

In September 2023, the chancellor established the Task Force on Generative Artificial Intelligence (AI). The charge to the task force was:

> *"I charge the Joint Task Force on Generative AI to examine the following:*
> *• current and future use cases of generative AI at educational institutions;*
> *• university processes that could consider utilizing generative AI technologies responsibly;*
> *• potential impacts of the appropriate adoption of generative AI on teaching, research, service, and administrative processes and its use by external vendors including an examination of both the appropriate and inappropriate use such technology; and*
> *• the implications of generative AI for educational records, research data, scholarship and creative works, rights in relation to university-based data (e.g., files, audio and video recordings), intellectual property, consent, and privacy.*
>
> *The Joint Task Force on Generative AI shall make recommendations related to the appropriate use of generative AI at UMass Amherst."*

This report summarizes the work of the task force that was conducted in the academic year 2023/24 and the recommendations that are presented to the faculty senate and to the administration as the outcome of this work.

II. Process

The Joint Task Force on Generative Artificial Intelligence (JTFGAI) convened from October 2023 until April 2024. The task force decided to divide into the following four subgroups to conduct more focused discussions on these topics:

- Education
- Research
- Operations and Compliance
- Privacy, Bias, and Ethics

The subgroups generated a series of recommendations for both the Faculty Senate and the University Administration to consider. These recommendations were developed through working group discussions in each area, discussed by the task-force as a whole, and then formally adopted.

The task force added the following individuals as non-voting members to provide additional expertise: Matthew Dalton, Brad Wheeler, Lisa Wolf, and Kate Woodmansee.

The task force provided an update of its work to the faculty senate at the December 13, 2023 meeting of the faculty senate. The task force also contributed to the UMass system-wide Symposium on Artificial Intelligence that was held on April 11, 2024 at the Mt. Ida campus of UMass Amherst.

The reports from the four subgroups are included in this report in Section V. These subgroup reports contain a more detailed discussion of the various issues that were considered.

III. Recommendations

The task force articulated the following underlying values:

1. The university values the exploration of new ideas and supports the thoughtful application of new technologies, such as generative artificial intelligence (GAI), to education, research, and administrative processes when appropriate.
2. The university values education and research endeavors as processes with human-to-human interaction at the core. Technological tools, such as GAI, provide support for but are not a substitute for these processes.
3. The university's mission and values must inform any policies or sanctioned practices involving GAI. The use of GAI should not reduce equitable access to a human-centered, effective, and quality educational experience or impinge on rights to academic freedom, intellectual freedom, privacy, due process, or any other substantive rights.

The task force makes the following policy recommendations:

1. The university shall develop and administer training for faculty, students, and staff about how GAI technology works, its limitations, and its potential uses.
2. The use of GAI shall be guided by balancing the potential positive impact of its use with the level of risk inherent to the technology. The use of GAI is not appropriate for some decisions and activities. In addition to university-wide use policies and prohibitions, units should develop, as appropriate, standards and guidelines for use and limitations of use for their operations. Standards and guidelines in academic units should be consistent with academic freedom and individual decision-making in instruction and research.
3. Any member of the university community who uses GAI shall be responsible for and accountable for any GAI-generated material as if they had created it themselves without GAI assistance.
4. Any member of the university community who uses GAI in their university role shall appropriately disclose such use. In addition, university applications that capture and store data from members of the university community shall employ principles of notice, including any data used in GAI model training or improvement.

5. Any member of the university community who uses GAI in their university work shall follow principles of consent for that use, as appropriate. University applications that capture and store data from members of the university community shall also employ principles of consent, as appropriate, including for any data used in GAI model training or improvement.
6. The university shall work with its vendors to ensure that principles of notice and consent are being followed by systems that employ GAI.
7. Any use of GAI must adhere to legal and university policies on data privacy and compliance, such as, but not limited to, academic integrity, the BOT Policy on Responsible Conduct of Research and Scholarly Activities, FERPA, HIPAA, and IRB protocols.

IV. Next Steps

We expect that the faculty senate will take up these recommendations and pursue next steps of their implementation. We recommend to the Rules Committee to consider referring the recommendations in this report to faculty senate councils as appropriate. These might include:

- Academic matters council
- General education council
- Graduate council
- Information and communication technology council
- Research council
- Undergraduate education council

We suggest that these councils consider if any of the policies in their purview need to be adapted in light of the recommendations in this report.

V. Subgroup Reports

The following four lists of values/recommendations are excerpted from the complete subgroup reports (which appear in their entirety in the remaining pages of this document).

(Education)

Our recommendations are created within the context of the following stated core values:

- Students have the right to learn in class, and uses of technology should facilitate rather than serve as a substitute for the activities needed to support student learning
- Faculty should be well-informed and empowered to choose the best technology policies for their classes, and should explicitly communicate these policies
- Fairness in the classroom is paramount, and access to technology should not serve as an advantage to students when demonstrating mastery

- Students and their instructors have a right to know when and how AI systems are being used in their classrooms as part of their teaching and learning experiences

Crafting a GAI Usage Policy

Instructors should decide exactly when and how GAI use is permitted for their courses and/or specific assignments. Instructors should specify clearly in their syllabus if GAI may be used, how and when it may be used, and with what degree of attribution. Instructors may state a policy that permits GAI use to varying degrees or instructors may choose to disallow GAI use for class assignments and activities. Regardless of the policy instructors select, it is important for instructors to recognize that with GAI readily available to students, there may need to be pedagogical modifications to the design of assessments and course deliverables. The nature of those pedagogical modifications will depend on the learning goals for a given course and the expectations of the instructor.

(**Research)**

Recommendations: Use of AI tools will soon be employed in most, if not all, domains of academic, artistic and creative, industrial, and government work. Therefore, creating space and resources on campus for education and training about the effective and responsible use of AI tools is paramount.

- *Recommendation #1: Build out the grid*. The Research Subgroup of the JTFGenAI recommends that an interdisciplinary group of faculty and staff, which may include some or all members of this subgroup, continue to work on filling in the grid with the considerations, expectations, responsibilities and opportunities related to the use of AI across all these phases of research from formulating research ideas and experimentation/data collection to the evaluation, reproducibility and dissemination of results. This work should include a content analysis to remove duplication and determine whether the content has been appropriately categorized as generally applicable to the use of AI across all domains of research and scholarship or specific to a particular broad field of study. The content in the grid may then be used to inform the development of educational offerings.
- *Recommendation #2: Develop educational programming.* The Research Subgroup recommends the development of educational programming on artificial intelligence of the following varieties: 1) undergraduate and graduate courses, 2) training modules (e.g. for new students, lab members, employees), and 3) guidelines for informal conversations between students and their advisors.

  Educational programming must necessarily equip members of our campus community to make use of powerful AI tools to advance their research and scholarship while simultaneously understanding and adhering to the guardrails that constrain the unethical

and potentially illegal use of AI in the conduct of research and scholarship. These tenets of capacity building and compliance should guide all educational programming. Moreover, it is also critical that educational programming be lightweight and adaptable to change in step with the rapidly evolving landscape of AI technology.The use of third-party, commercial AI tools, such as the Large Language Models (LLMs) developed by OpenAI, Microsoft, and Google, poses a security and privacy challenges to research where protection of intellectual property and the ethical treatment of human subjects, including securing their personal and confidential information, is a top priority.

- *Recommendation #3: Provide guidance on the use of LLMs in research.* The Research Subgroup recommends that a team of experts from the Libraries, IT, Research & Engagement, provide guidance to the campus research community, perhaps in the form of a webpage or SharePoint site, specifically on the use LLMs to ensure that research security, privacy and confidentiality are at the level of best practice and in alignment with professional community guidelines. The recommended guidance should prepare researchers to ascertain and follow the lead of funding agencies, professional communities and publications about what is considered acceptable use of AI in their research.
- *Recommendation #4: Survey the use of AI in research.* The Research Subgroup of the JTFGenAI recommends that the campus endeavor to understand where and how AI is used in the process of research using multiple methods, including at least a campuswide survey. This environmental scan will inform the development of guidance on the use of AI recommended above, which necessarily includes the expectation that disclosure of the use of AI tools in research is required to ensure transparency of methods.
- *Recommendation #5: Employ an institutional review process for the use of AI in human-subjects research.* The Research Subgroup of the JTFGenAI recommends that deployment of an IRB, or IRB-like processes, for the use of AI in human subjects research that is guided by national best practices as described by leading national organizations like PRIM&R.

(Operations and Compliance)

Many data protection regulations exist that impact the use of AI, including FERPA, HIPAA, GLBA, and in some instances, GDPR. To best address these disparate regulations, UMass Amherst should adopt a standards-based and principled approach. A 2020 Harvard research paper[1] reviewed 36 “principles” documents and found that there are 8 main principles and 47 subprinciples that need to be addressed when using AI:

1. Privacy
2. Accountability
3. Safety and Security
4. Transparency and Explainability

5. Fairness and Non-discrimination
6. Human Control of Technology
7. Professional Responsibility
8. Promotion of Human Values

A helpful diagram showing the correlation of current laws and guidlines with these principles is available at: http://wilkins.law.harvard.edu/misc/PrincipledAI_FinalGraphic.jpg. For the complete list of principles and sub-principles, see Appendix A.

UMass Amherst must focus on four main areas to effectively address these principles:

1. In order to ensure that external companies and partners align with UMass Amherst's legal and ethical obligations, UMass must address these compliance and ethical concerns during the **procurement** process with all vendors.
2. In order to ensure that the application of Information Technology Policies by the UMass Amherst community aligns with the principles, UMass must provide appropriate **standards and guidelines** for how to interpret the Information Policies in various contexts when using Generative and other types of AI.
3. In order to ensure proper understanding by the UMass Amherst community, UMass must provide **Awareness and training** in a variety of forms on the appropriate use of AI in the major contexts of the University.
4. In order to ensure that UMass Amherst maintains a prominent position with regards to the legal and ethical use of IA, UMass must foster strong **community engagement** in the larger higher education community as these issues evolve. UMass must provide thought leadership where appropriate and be part of the conversations where they happen to further enrich our own stance and the larger community narrative.

(Privacy, Bias, and Ethics)

Recommendations For Umass:

1. **UMass must provide clear guidelines on sharing data with AI agents. In particular, the campus must identify types of data that <u>may not</u> be shared with general access AI agents. As with other IT tools, the campus should negotiate with vendors to ensure the protection of and ethical and legal uses of data.**
   The list of legal and ethical concerns with AI is lengthy – privacy, export control law, creators' legal and moral rights, research rights, contractual rights, human subject interests, etc. – and the relevant concerns will vary widely depending on functions within the university. Consequently, this will require careful assessment in developing contracting practices in procurement, and in other key areas – research, student assessment, awards, etc.
   It will also require a robust education initiative available to all staff, since staff at all

levels of the university may encounter AI agents on the Internet, as well as embedded in ordinary software and apps. Two priorities include:

- FERPA: All staff who work with student records should receive guidance around FERPA and AI agents, including understanding that using AI agents implicates FERPA concerns.
- All staff should understand that AI agents available on the web or in other software applications may incorporate data into their models, and thus sharing data – i.e., copy/pasting, uploading, or otherwise importing / ingesting data into AI agents – must be assessed carefully under the same rubrics applicable to sharing data with any other third party.

2. **UMass must provide clear guidance as to the use of AI in decision-making, evaluation, and assessment contexts. Use of AI in particular high-stakes or sensitive decision-making contexts should be prohibited.**
   For example, AI tools should never be used to assess the quality of scholarship or student work, to assess merit for financial aid, or to compare candidates for jobs, among other high-impact or sensitive areas. Such uses present insuperable concerns about bias, as well as accuracy, accountability, and transparency / openness. Moreover, even if there is no apparent or obvious effect, there may be less-obvious or second-order bias effects. Some clear examples include (but are not limited to):
   - Student related actions and assessments, including admissions, financial aid, instructional activities such as grading coursework, and advising
   - Personnel actions, such as hiring, promotion, tenure, assessment, and discipline
   - Assessments for potential disciplinary or criminal scrutiny

   This guidance also requires a robust educational initiative available to all staff, to understand how and when – and when not – to use AI tools in analysis or other decision making support contexts. This must include the risks relating to bias, as well as to larger concerns relating to accuracy and reliability.

3. **UMass must provide clear guidance on the use of interactive AI or AI-generated works in communications.** For example, the use of AI to write a letter of recommendation or an expression of grief are always inappropriate, because it degrades the human connection that such circumstances demand. Similarly, while the use of AI to conduct interactive conversations may be appropriate in low-level phone trees, it would be inappropriate in situations in which a person needs or expects empathy, human judgment, or connection.

4. **The use of AI in any context – communication, generative, analytical, or otherwise – should be guided by principles of transparency.** The University must develop policy to make this clear. While this principle is clearly drawn from academic norms and therefore

intuitive in academic contexts, it is also important in non-academic contexts – such as interactive communications, official communications from the university, and the wide array of business operations the University conducts.

5. **Education in Balancing Risks: The University must develop educational programs to support staff in understanding and using AI appropriately.**
   Beyond the clear guidelines that we recommended to be set forth above, we also note that **any** use of AI should be carefully assessed, balancing the potential benefits of that use with the potential or known risks. To do so, staff need education about AI – how it works and doesn't work, what the risks and concerns are – in order to appropriately assess whether or not to use it in a given situation. The University has a duty to provide the resources to staff to understand how AI tools function, and the risks of bias and other inaccuracies, privacy and data security, and the larger ethical and social concerns that AI raises. AI functions have increasingly been added to other workforce tools – including transcription software, word processing, and search engines, to name a few. All staff – administrative, academic, faculty, graduate employees – are already encountering AI tools. They need the resources to assess the risks, understand the potential benefits, and can therefore appropriately balance the costs and benefits for use in a given situation.

6. Finally, we note that the possible impact of AI poses both a **duty to educate our students** and an **opportunity to bring UMass' values to this area.**
   UMass is a leader among universities in sustainability, social justice, and ethical concerns, and in particular its expertise in critical media / information studies and public interests in technology. We are well-positioned to establish UMass as a leader in assessing these concerns – and, it is our opinion that this is one of the ways we will best serve our students

COMPLETE SUBGROUP REPORTS

a. Report by the Subgroup on Education

This guide was created by the Education subcommittee of the Joint Task Force on Generative Artificial Intelligence, made up of Matthew Rattigan (coordinator), Marc Liberatore, Laura Quilter, Farshid Hajir, Claire Hamilton, Michelle Trim, Anthony Paik, Allison Butler

Executive summary:

Modern, publicly available Generative Artificial Intelligence tools have progressed to the point where they can mimic different types of student work products. In order to ensure that learning goals are effectively met, instructors need to understand how GAI

use might subvert intended learning activities, and craft policies to suit their individual subject expertise and pedagogical approach.

Introduction

Generative Artificial Intelligence (GAI) is a category of machine-learning-based techniques that can produce text, images, and computer code in response to natural language prompts. In recent years, several Large Language Models (LLMs) have been deployed publicly, allowing users to anonymously submit requests and receive output. As the quality of the underlying models has improved, many popular LLMs have become capable of mimicking human intellectual products (including speech, writing, and creative works) in response to prompts based on common tasks assigned as part of an academic course (e.g., quiz questions, programming tasks, writing prompts). While the core student learning objectives in a university setting are unchanged by the availability of GAI, the methods of how instructors accomplish certain pedagogical goals might. The goal of this document is to provide the Joint Task Force on Generative AI with recommendations for policies, aids, and ongoing structures that help delineate responsibility. Some of the supports we suggest could provide instructors with an understanding of how student use (or misuse) of common GAI tools might affect students' foundational learning, some course designs, and some common assessments. Where appropriate, we provide guidance and recommendations for course design and pedagogy, with the understanding that instructors' goals within their individual contexts will vary. In addition, we hope the insights from this document will inform administration with strategic guidance in the crafting of university policy.

Committee and Values

Our recommendations are created within the context of the following stated core values:

- Students have the right to learn in class, and uses of technology should facilitate rather than serve as a substitute for the activities needed to support student learning
- Faculty should be well-informed and empowered to choose the best technology policies for their classes, and should explicitly communicate these policies
- Fairness in the classroom is paramount, and access to technology should not serve as an advantage to students when demonstrating mastery
- Students and their instructors have a right to know when and how AI systems are being used in their classrooms as part of their teaching and learning experiences

Crafting a GAI Usage Policy

Instructors should decide exactly when and how GAI use is permitted for their courses and/or specific assignments. Instructors should specify clearly in their syllabus if GAI may be used, how and when it may be used, and with what degree of attribution.

Instructors may state a policy that permits GAI use to varying degrees or instructors may choose to disallow GAI use for class assignments and activities. Regardless of the policy instructors select, it is important for instructors to recognize that with GAI readily available to students, there may need to be pedagogical modifications to the design of assessments and course deliverables. The nature of those pedagogical modifications will depend on the learning goals for a given course and the expectations of the instructor.

Common Issues Related to GAI, Teaching, and Learning

It is important that the university community be made aware that students can privately, and often undetectably, utilize GAI when completing coursework. While in some contexts, the use of GAI tools may complement or even enhance the learning process, in others, it may serve as a substitute for learning and weaken some pedagogical strategies. Furthermore, student use of GAI tools might severely impact the validity of certain forms of assessment. Below we outline some common types of classroom tasks and assignments, and outline how GAI may affect their value and intended purpose.

GAI for Writing in English

GAI tools such as ChatGPT are capable of producing text that mimics speech common to university contexts. This GAI text is often flawed (for example, it might include missing, invalid, or inappropriate citations), it cannot effectively use context to develop or structure ideas, it will not be able to generate ideas or facts related to world events that have occurred since the model was trained (usually a year or two in the past), it does not actually understand the language that it is producing and thus can create complex structured sentences that seem advanced, but often convey simplistic ideas when analyzed, and (as a result) is at-present usually detectable by a human expert. While tools such as Turnitin claim to have highly accurate GAI-detection capabilities, experts have found all electronic detection methods to be unreliable. As we move into the future, it is possible that GAI systems will produce text that is a closer approximation of human-produced text, and it may become increasingly difficult for instructors to differentiate between GAI and student-created text.

A student could choose to use GAI at any stage of the writing process. For instance, a student might use a GAI tool to produce a critical essay whole cloth from an assignment prompt. Alternatively, they might use GAI to produce a detailed outline of an essay based on parameters from the assignment and then use that outline to produce text

themselves. Another student might only employ GAI to check for grammatical errors or usage fluency, which can seem to work well or can in some cases be risky since there is no guarantee that GAI is more correct in usage, particularly for disciplinary contexts. In general, and similar to other forms of assisted writing, the closer to the beginning of the writing process that a student uses GAI, and the more actual text that they produce themselves, the more difficult it is to determine whether a GAI tool has been used at an early stage based on the final product.

Given the different ways GAI can be employed, instructors should be cautious about relying on a single writing product, such as a final draft, to assess student learning outcomes. Since the 1980s, the literature of writing pedagogy has advocated for a process-based rather than product-based approach to the assessment of learning outcomes in writing. A single work product outside of writing courses might be used as evidence of achievement of higher-level learning goals (e.g., awareness of disciplinary conventions, understanding of course content, appropriate uses of evidence); some uses of GAI tools may weaken the integrity of these kinds of assessments. This means that instructors may want to reconsider relying exclusively on final written products as demonstrations of understanding of course concepts, particularly concepts common to foundational courses in a discipline or common general education courses. For example, if illustrating an understanding of the way art styles change over time and space is an important learning goal for an introductory general education course, then instructors need to design assessments that examine students' thinking on that topic from a process-based set of scaffolded activities/assessments and perhaps not take on faith that a finished essay is proof that higher-level thinking or synthesis of course material took place during its creation. If instructors wish to discourage or disincentivize the use of GAI in a class, writing prompts that require engagement with specific lectures, readings, examples from class, and other context specific ideas will be more difficult (and likely less worthwhile) for students to accomplish using GAI. Regardless of the permissiveness instructors articulate for using GAI in their courses, it is imperative that the university inform instructors as to the affordances and limitations of GAI so that they can design their assessments with the availability of these tools in mind.

GAI Scholarly Reading and Discussion

Many undergraduate and graduate seminars are based on a practice of shared reading, presentation, and group discussion. Often, the ability to read, process, and synthesize relevant information across scholarly works is a desired learning outcome in itself, in addition to increasing area-specific knowledge. Some instructors traditionally rely on participation in discussions to determine whether students are keeping up with and understanding assigned reading. However, GAI tools can also be used to "summarize" large amounts of text. Of course, these summaries may be incomplete, or even factually

inaccurate (though, those weaknesses could improve in the future). If students utilize GAI tools to reduce their reading load by reducing longform works into bullet points, they may circumvent some of the critical thinking typically assumed to have been part of the process, but may still be able to convincingly participate in seminar discussions.

GAI for Computer Code Generation

Traditionally, the ability to produce functional and correct code has served as a tangible marker of comprehension and mastery of underlying principles, algorithms, and concepts in computer science and other STEM courses that rely in part on computer programming. As a result, creating code has been a means for students to demonstrate understanding of course concepts (and of programming itself). However, modern GAI code generation tools have reached the point where they are able to effectively mimic the type of code that would otherwise demonstrate mastery of higher level learning in technical areas.

While code generation tools have long existed to aid in producing scaffolding and boilerplate code, they have typically required substantial developer input and a solid grasp of programming fundamentals. However, the application of GAI in code generation has revolutionized this landscape. Now, these models possess the capability to generate entire programs from natural language instructions, often yielding code that is indistinguishable from that produced by human programmers. This ability is especially true in the domain of the kinds of programs often required in introductory programming courses, where an extraordinary amount of training material exists and has been ingested by these models – their facility in this domain is extremely strong.

Yet, amidst this technological advancement, it is imperative to recognize that the role of generative AI should not supplant the need for instruction in coding and algorithmic thinking. Just as calculators haven't rendered arithmetic instruction obsolete, GAI should complement rather than replace traditional pedagogical approaches. Fundamental to this stance is the understanding that proficiency in basic computing and programming concepts remains indispensable. While generative AI can streamline certain aspects of coding, it does not negate the necessity for learners to cultivate a robust foundation in computational thinking. Mastery of these fundamentals serves as a springboard for fostering higher-order cognitive skills, such as problem decomposition, algorithmic reasoning, and the like.

This analogy between calculators and generative AI is useful: Just as most individuals employ calculators in their daily computations, it is conceivable that programmers may utilize GAI tools in their workflow. However, akin to how calculator usage does not obviate the need for understanding arithmetic principles, reliance on generative AI necessitates a comprehension of underlying concepts. In addition, calculators are reliable

producers of correct answers, assuming they are asked the right question. There is no guarantee that questions about mathematical or coding problems will always receive correct answers from GAI tools. Employing GAI without a solid grasp of programming fundamentals risks divorcing practitioners from the intrinsic logic and rationale behind the code they produce, potentially impeding their ability to troubleshoot, innovate, and adapt to new problems.

Creative Works

In addition to writing and code, many publicly available GAI tools can be used to generate creative works from a natural language text prompt, including images, audio, music, and video. As with generated text, the outputs can be flawed (e.g., synthetic photos of humans containing three arms, or videos of birds flying backwards). It should be noted that progress in these media has been rapid in the past several years, and future versions of GAI systems might improve to the point where there is no human-perceptible difference between GAI works and human ones.

As in the cases above, GAI might be deployed at any point in the creative process, from conception to execution (though in most cases the latter might be impossible for a given medium). A music student might use GAI to transcribe a piece of music to a different key. And while current LLMs are unable to produce non-digital works (LLMs can't paint… yet), a student might use an image generator to produce a sketch on which to base a tangible work with a certain theme or reflecting a specific style.

Considerations for Course Pedagogy

In all of the above cases, GAI might be used at several different points in the learning process. Instructors should be deliberate about which aspects of each activity or assignment are most important and are being assessed by any given product, and consider the following questions when designing assignments and assessments:

- What are the most important learning outcomes for the activity? Will these be achieved if GAI is used?
- How effective is a given assessment if students are able to use GAI to complete it?
- Do students understand which specific activities and practices are the most important for a task?
- Does the use of GAI enhance or improve any of the current course assignments or class activities?

Depending on their answers to the above questions, instructors should make deliberate choices about whether GAI use should be permitted for different course activities. In

some cases, allowing GAI use for completing assignments will clearly subvert student learning – for instance, permitting students in an introductory language course to use a GAI translator to complete practice exercises. For other activities, whether GAI use is appropriate depends heavily on the pedagogical context. Instructors should communicate their reasoning behind the policies they impose, and how it relates to learning goals for the course and methods of assessment, so students are able to understand how unauthorized GAI use might hinder their long term progress.

Some instructors may prefer to increase the use of in-class assessments and downweight the grade value of work completed outside of the classroom. While this solves the problem of unauthorized GAI use, it requires redesigning course activities. Instructors may want to avoid fostering a culture of distrust or presumed dishonesty. Like many classroom policies, GAI policies will require an element of trust between students and instructors. The level and dimensions of that trust will vary with context, and instructors should be explicit about their expectations for student behavior. Ultimately, we believe that decisions about GAI use for course activities must be the responsibility of individual instructors. Choices should be made deliberately, based on subject expertise, pedagogical approach, and a coherent understanding of how GAI might be used (or misused) in a given learning context, keeping in mind the inherent trade-offs related to different decisions and options.

b. Report by the Subgroup on Research

Membership of Research Subgroup: Francine Berman, Steve Bischof, Jeffrey Blanchard, Jennifer Normanly, Anthony Paik, Andre Kenneth Chase Randall, Michelle Trim, Jacqueline Urla, Loren Walker (coordinator), Lisa Wolf, Tilman Wolf, Ethan Zuckerman, Johanna Frances Yunker

Summary of Activities: The Research Subgroup of the Joint Task Force on Generative Artificial Intelligence (JTFGenAI) met via Zoom four times since November. Each meeting was recorded. In this report, we summarize the activities of the subgroup, present our preliminary findings, and offer several recommendations for next steps.

While most of the conversation thus far in the broader JTFGenAI has understandably been focused on the use of generative artificial intelligence in the classroom, there are of course significant implications for research. GenAI, and Artificial Intelligence tools more broadly, is expected to intersect with and impact research and scholarship across all the phases of the process, from formulating research ideas and experimentation/data collection to the evaluation, reproducibility and dissemination of results. The opportunities for the advancement of research and scholarship are substantial, as are the potential pitfalls.

Therefore, the research subgroup of the JTFGenAI has organized its thinking around the considerations, expectations, responsibilities *and* the opportunities related to the use of AI across all these phases of research. The group created a spreadsheet/grid for this purpose. The "Overall" tab is for ideas on the AI considerations, expectations, responsibilities and opportunities that could apply to all disciplines. The grid also contains additional tabs representing several broad disciplinary areas, based on the "broad field" categories used by Academic Analytics.

Two questions posed during our December meeting motivate this organizational approach: 1) What can this group do to inform the development of courses/trainings related to AI? 2) What can this group do to inform how "guardrail" frameworks, like IRBs, might apply to the use of AI in research. To explore the second question, the group invited Professor Lynnette Leidy Sievert, Chair of the campus IRB and Iris Jenkins, Associate Director of the Human Research Protection Office to share their perspectives on the IRB implications of AI during our meeting in January. Professor Jesse Mager, Interim Asst Vice Chancellor of Research & Engagement also participated. Iris Jenkins provided a readout from the national PRIM&R meetings for IRB personnel that she recently attended where AI was a major topic of discussion. Our January meeting with the IRB personnel focused on the intersections between the use of AI in research and the necessity of protecting human subjects and their data. It also became clear that human-subjects protection isn't the only lens through which this subgroup should consider recommendations for education/training and guardrails pertaining to the use of AI in research. Other considerations include: energy use/environmental impacts, data security, intellectual property ownership, and research ethics more generally.

Preliminary Findings: First, the research subgroup of the JTFGenAI asserts that the same principles of transparency, honesty, accountability, and privacy/confidentiality apply to the use of AI as would apply to any tools and processes used in the responsible conduct of research (RCR) and scholarship. These RCR norms, as they apply to AI, will be informed by an individual's academic community and often dictated by the editorial boards of the relevant peer-reviewed journals. Therefore, the "conversation" about acceptable, responsible, and ethical use of AI must be between the researcher/scholar and their community or, in the case of students, their advisor, not the university administration. We also hasten to point out that what's considered acceptable, responsible, and ethical use of AI will also vary by the type of work product, e.g. the standards for a dissertation may be quite different from the standards for a grant proposal.

Second, regarding the establishment of "guardrails" frameworks for the use of AI in research. Our research subgroup focused primarily on the protection of human subjects in research. There are national groups, including PRIM&R, who are already working on these issues as they pertain to the use of AI. Our faculty and staff working in the human subjects/IRB space are monitoring those developments. However, it should

also be acknowledged that there are other relevant considerations as well, including energy use/environmental impacts, data security, intellectual property ownership, and research ethics more generally.

Recommendations: Use of AI tools will soon be employed in most, if not all, domains of academic, artistic and creative, industrial, and government work. Therefore, creating space and resources on campus for education and training about the effective and responsible use of AI tools is paramount.

- *Recommendation #1: Build out the grid*. The Research Subgroup of the JTFGenAI recommends that an interdisciplinary group of faculty and staff, which may include some or all members of this subgroup, continue to work on filling in the grid with the considerations, expectations, responsibilities and opportunities related to the use of AI across all these phases of research from formulating research ideas and experimentation/data collection to the evaluation, reproducibility and dissemination of results. This work should include a content analysis to remove duplication and determine whether the content has been appropriately categorized as generally applicable to the use of AI across all domains of research and scholarship or specific to a particular broad field of study. The content in the grid may then be used to inform the development of educational offerings.
- *Recommendation #2: Develop educational programming.* The Research Subgroup recommends the development of educational programming on artificial intelligence of the following varieties: 1) undergraduate and graduate courses, 2) training modules (e.g. for new students, lab members, employees), and 3) guidelines for informal conversations between students and their advisors.

  Educational programming must necessarily equip members of our campus community to make use of powerful AI tools to advance their research and scholarship while simultaneously understanding and adhering to the guardrails that constrain the unethical and potentially illegal use of AI in the conduct of research and scholarship. These tenets of capacity building and compliance should guide all educational programming. Moreover, it is also critical that educational programming be lightweight and adaptable to change in step with the rapidly evolving landscape of AI technology.

  The use of third-party, commercial AI tools, such as the Large Language Models (LLMs) developed by OpenAI, Microsoft, and Google, poses a security and privacy challenges to research where protection of intellectual property and the ethical treatment of human subjects, including securing their personal and confidential information, is a top priority.

- *Recommendation #3: Provide guidance on the use of LLMs in research.* The Research Subgroup recommends that a team of experts from the Libraries, IT, Research & Engagement, provide guidance to the campus research community, perhaps in the form of a webpage or SharePoint site, specifically on the use LLMs to ensure that research security, privacy and confidentiality are at the level of best practice and in alignment with professional community guidelines. The recommended guidance should prepare researchers to ascertain and follow the lead of funding agencies, professional communities and publications about what is considered acceptable use of AI in their research.
- *Recommendation #4: Survey the use of AI in research.* The Research Subgroup of the JTFGenAI recommends that the campus endeavor to understand where and how AI is used in the process of research using multiple methods, including at least a campus-wide survey. This environmental scan will inform the development of guidance on the use of AI recommended above, which necessarily includes the expectation that disclosure of the use of AI tools in research is required to ensure transparency of methods.
- *Recommendation #5: Employ an institutional review process for the use of AI in human-subjects research.* The Research Subgroup of the JTFGenAI recommends that deployment of an IRB, or IRB-like processes, for the use of AI in human subjects research that is guided by national best practices as described by leading national organizations like PRIM&R.

c. Report by the Subgroup on Operations and Compliance

Membership of Operations and Compliance Subgroup: Steve Bischof, Matthew Dalton (co-coordinator), Laura Haas, Kevin Kerwood, Chris Misra (co-coordinator), William Murphy, Anthony Paik, Erin Schadel, Michelle Trim, Brad Wheeler, Tilman Wolf, and Kate Woodmansee

Introduction:

While the benefits of artificial intelligence can be transformational, there are compliance issues that any institution must address regarding legal and ethical issues. There is no overarching law federally that articulates how artificial intelligence should be used, and there is a patchwork of regulations at the state level, and even more fragmentation when it comes to state regulations. However, a consensus of sorts has emerged. Many data protection regulations exist that impact the use of AI, including FERPA, HIPAA, GLBA, and in some instances, GDPR. To best address these disparate regulations, UMass Amherst should adopt a standards-based and principled approach. A 2020 Harvard

research paper[1] reviewed 36 “principles” documents and found that there are 8 main principles and 47 subprinciples that need to be addressed when using AI:

1. Privacy
2. Accountability
3. Safety and Security
4. Transparency and Explainability
5. Fairness and Non-discrimination
6. Human Control of Technology
7. Professional Responsibility
8. Promotion of Human Values

A helpful diagram showing the correlation of current laws and guidlines with these principles is available at: http://wilkins.law.harvard.edu/misc/PrincipledAI_FinalGraphic.jpg. For the complete list of principles and sub-principles, see Appendix A.

UMass Amherst must focus on four main areas to effectively address these principles:

2. In order to ensure that external companies and partners align with UMass Amherst's legal and ethical obligations, UMass must address these compliance and ethical concerns during the **procurement** process with all vendors.
3. In order to ensure that the application of Information Technology Policies by the UMass Amherst community aligns with the principles, UMass must provide appropriate **standards and guidelines** for how to interpret the Information Policies in various contexts when using Generative and other types of AI.
4. In order to ensure proper understanding by the UMass Amherst community, UMass must provide **Awareness and training** in a variety of forms on the appropriate use of AI in the major contexts of the University.
5. In order to ensure that UMass Amherst maintains a prominent position with regards to the legal and ethical use of IA, UMass must foster strong **community engagement** in the larger higher education community as these issues evolve. UMass must provide thought leadership where appropriate and be part of the conversations where they happen to further enrich our own stance and the larger community narrative.

Procurement

Working with UPST and with others, the University should develop standard contract language as part of our terms and conditions so that any existing and future contracts as they relate to AI better align with these principles. In the procurement and contracting process for AI systems, it is crucial to establish clear expectations and protections when it comes to the principles articulated above. Any AI system being considered should have a thorough risk assessment conducted against these principles as part of the evaluation

process. Contracts should mandate regular audits and updates to AI algorithms by the company to ensure continued alignment with the principles.

Furthermore, to minimize the environmental impact of AI, contracts should stipulate the use of energy-efficient hardware and algorithms. Developers can be encouraged to design AI systems that reduce computational redundancy and promote longevity, thereby decreasing electronic waste. The procurement process should also prioritize vendors who demonstrate a commitment to sustainable practices, including the use of renewable energy sources and adherence to environmental standards. By taking these steps, organizations can ensure that their use of AI supports not only ethical data practices but also environmental stewardship.

Standards

UMass Amherst must establish clear standards and guidelines to interpret the Information Technology Policies, including the Acceptable Use Policy, Information Security Policy, and Information Privacy Policy as they relate to the appropriate use of AI. These guidelines should articulate the principles, and how they are to be applied in a variety of contexts, including the instructional, academic, research, and administrative activities of the University. These standards and guidelines should be clear enough that they provide meaningful guidance, but flexible enough that they maintain relevance as the various uses of AI evolve.

Education

As UMass Amherst formulates its standards and guidelines, it is important to concurrently craft training and awareness materials for AI literacy. This will enhance the community's understanding and adoption of these principles across various campus activities. These should be developed to be accessible and relevant to the variety of individuals and applications of AI that are found within the community. The AI literacy materials should be evaluated regularly for effectiveness and adapted to the changing AI landscape. For similar reasons, these updated materials should be continuously presented to the community.

In addition to these AI literacy materials, UMass Amherst must engage in internal dialogue through the faculty, advisory committees, focus groups, and other venues in order to ensure that fresh understandings and voices with regards to AI are regularly encouraged and incorporated into other efforts.

Community Engagement

Community engagement is vital for the continued understanding of the current and future implications of artificial intelligence for UMass Amherst, and the ability to influence state and national policy. Higher education technology organizations such as

EDUCAUSE, Internet2, CNI, and REN-ISAC are critical for UMass Amherst to engage in meaningful conversations with regards to artificial intelligence at this national level, as are non-technical organizations such as AAUA[2], AAUP[3], NACUBO[4], AIR[5], and AGB[6]. These institutions allow UMass to provide thought leadership, and to understand how others are looking at the impact of AI.

In addition to engaging in collaboration with professional organizations in higher education, to maintain relevance in the evolution and policy of AI, UMass Amherst should regularly engage with key partners in industry, whether by serving on advisory boards, or by regular executive sessions with company leadership. It should also engage with state policy makers and regional leaders to influence the development of new policies and regulations in ways that align with the principles to which UMass adheres.

Conclusion:

It is only through a coordinated, comprehensive approach that UMass Amherst can align its partnerships and its community to the principles that are essential for the legal and ethical use of AI, and therefore to meet its compliance obligations. By engaging in the process outlined above, UMass positions itself to not only be able to address the compliance obligations of today but ensure continued alignment and improvements for the future.

d. Report by the Subgroup on Privacy, Bias, and Ethics

Membership of Privacy, Bias, and Ethics Subgroup: Allison Butler, Matthew Dalton, Claudia Donald, Anthony Paik, Laura Quilter (coordinator), Matthew Rattigan, Michelle Trim, Ruth-Ellen Verock-O'Loughlin, Tilman Wolf, and Ethan Zuckerman

The Privacy, Bias, and Ethics subgroup has identified five key principles relating to use of (generative) artificial intelligence tools on campus, and made six specific recommendations of areas for development of policy and procedure. Ethan Zuckerman is also supervising a research project aimed at surveying and summarizing ongoing privacy, bias, and other ethical concerns in the AI sphere.

Generative artificial intelligence ("AI") tools present many ethical concerns to our campus community. Privacy and bias are foremost among them, but other ethical concerns include the degradation of human communications; the consent of those whose communications, data, or writings may be assessed by or incorporated into AI models; the energy use and environmental concerns of AI models; the workforce concerns, including the fairness of labor used in developing AI tools, as well as potential workforce displacement – to name just a few of the concerns that have been raised by scholars,

activists, whistleblowers, and many others. We are certain that as AI technologies mature and evolve, additional concerns will arise.

This report, therefore, is necessarily only a first pass – it will be important to revisit these issues as AI technologies and applications evolve. This report is also, necessarily, a high-level perspective, aimed at describing principles. While we provide examples, particularly aimed at sketching the clear poles, and identifying gray areas, these principles will need to be carefully applied to specific scenarios, departments, workflows, use cases, and so on. There are numerous circumstances in which the use of AI – whether to analyze existing data or create a new work – is illegal, unethical, inappropriate, or otherwise ill-advised.

Our principles are drawn from the values of UMass – as a public, land grant established, educational and research institution, with a public service mission, that prioritizes excellence; diversity, equity, and inclusiveness; freedom and openness; integrity, stewardship, and sustainability; innovation and creativity; and impactful, positive contributions to the welfare of people locally and globally.[1] These values inform all our endeavors, as educators, scholars and researchers, and administrators and staff running this institution. These values also inform the uses of AI by members of this campus community.

Based on the UMass values articulated above, and drawing from Fair Information Practice Principles and other guidance, our subgroup developed the following principles for uses of AI. These principles complement and augment other data protection and ethical practices already in place, from the Acceptable Use policy, to Human Subjects Research protocols, etc.

PRINCIPLES FOR USING AI

- **Transparency & Disclosure: Uses of AI should be transparent and disclosed.** Scholarly, academic, educational, and ethical norms all suggest that the use of AI should be disclosed. Legal regimes aimed at ensuring rights of persons affected by AI decision-making or analysis also indicate transparency. Transparency norms should apply to uses of AI including (i) generating new works (text, images, video, music, software code, etc.), or ideas, (ii) analyzing, summarizing, organizing, translating, etc., data that is provided to it; (iii) analyzing, summarizing, organizing, translating, etc., existing data on the Internet or in its

[1] UMass Strategic Plan, 2018-2023

knowledge base; (iv) communicating in live interactions.

Examples:

- ChatGPT is used to generate a paragraph of review material for a paper, class, or report. This should be cited.
- An AI agent such as a chatbot is used in an interactive conversation to triage requests for assistance. The AI agent should be clearly marked as such, so the user is not confused into thinking they are speaking with a human.

- **Fairness & Bias: AI should not be used to perpetuate or amplify existing biases present in society.**
  The research is clear that AI tools, which are developed using existing cultural products (created works, interactions, etc.) exhibit the same biases. Such biases are both inaccurate, and unjust. Use of AI therefore can and does perpetuate or even amplify those existing biases. This is unacceptable.

  Moreover, use of AI in decision-making poses a second-level difficulty, in that it may not be possible to replicate, assess, or understand the tool's decision-making. Thus, the use of AI in decision-making runs the risk not just of bias, but of *disguised* bias – violating the principles of transparency and disclosure above, and accountability and responsibility below.

  Therefore, AI tools should be barred altogether in cases where decisions are particularly material or high stakes, and even in low-stakes cases where biases are likely to affect the output. This includes all assessments of student work, employment performance and job candidates, quality of scholarship, any disciplinary processes or awards.

  Examples:

  - AI tools should be barred from assessing and ranking job candidate's writing style, because this is high stakes.
  - AI tools should be barred from use in providing feedback to students on their writing, even ungraded feedback, because AI tools have been shown to have cultural bias in evaluating writing.
  - AI tools should not be used as a singular "check for bias", because AI tools have themselves been shown to be biased. While this is a promising area of potential application for AI tools, at present, it is not reliable, and poses the risk of reifying the bias in a decision-making process.

In situations where the stakes are low, we suggest a balancing test: **the risk of a potentially erroneous or biased output of an AI should be commensurate to the stakes of a task or decision.**

- **Accountability and Responsibility: Humans who use AI are responsible for their work product and decision-making, regardless of any contributions by AI.**
  Due to the "black box" aspect of AI – the inability of users to assess or replicate AI assessments – requires that AI use must be transparent and disclosed to affected parties, to ensure accountability and responsibility.

  Example:
  - An author uses an AI to fact-check their text. It is later discovered that the facts that were checked were wrongly deemed correct. The author is responsible for the factual accuracy of their work, including fact-checking; using an AI system does not absolve the author of their responsibility or excuse any errors relating to faulty fact-checking.

- **Consent of participants: Individuals should be given the opportunity to choose whether or not to interact with, be assessed by, or contribute content to AI systems.**
  Due to the numerous significant ethical considerations that AI tools currently present – from environmental costs, labor and workforce concerns, to privacy, bias, and copyright infringement, just to name a few – individuals may reasonably wish to opt out of contributions to or interactions with AI agents.

  Examples of how to implement this:
  - The University, in negotiating access to AI tools, should require meaningful user consent protocols. In some cases, opt-in might be appropriate, and in others, notice and opt-out might be appropriate. In deciding the appropriate user consent protocol, the emphasis should be on empowering the individual to control their own data.
  - An instructor wishes to have students submit their work to AI agents for review / commentary. Students should be able to opt out of this assignment. If it is a key part of a class, then students should know that in advance of signing up to the class.

- **Redressability / Human review: Individuals should be given the opportunity to request human review of decisions made by or with AI input.**
  Best practices for data collections and algorithm-based decision making have long

included opportunities for redress for errors or suspected errors, including human review of data and decision-making.[2] This is particularly important given the inherent "black box" nature of AI analyses.

RECOMMENDATIONS FOR UMASS:

We add these specific recommendations, aimed at implementing the principles discussed above.

7. **UMass must provide clear guidelines on sharing data with AI agents. In particular, the campus must identify types of data that <u>may not</u> be shared with general access AI agents. As with other IT tools, the campus should negotiate with vendors to ensure the protection of and ethical and legal uses of data.**

   The list of legal and ethical concerns with AI is lengthy – privacy, export control law, creators' legal and moral rights, research rights, contractual rights, human subject interests, etc. – and the relevant concerns will vary widely depending on functions within the university.

   Consequently, this will require careful assessment in developing contracting practices in procurement, and in other key areas – research, student assessment, awards, etc.

   It will also require a robust education initiative available to all staff, since staff at all levels of the university may encounter AI agents on the Internet, as well as embedded in ordinary software and apps.

   Two priorities include:

   - FERPA: All staff who work with student records should receive guidance around FERPA and AI agents, including understanding that using AI agents implicates FERPA concerns.
   - All staff should understand that AI agents available on the web or in other software applications may incorporate data into their models, and thus sharing data – i.e., copy/pasting, uploading, or otherwise importing / ingesting data into AI agents – must be assessed carefully under the same rubrics applicable to sharing data with any other third party.

[2] Beginning with the introduction of Fair Information Practice Principles (FIPPs) in 1973 by the U.S. Department of Health, Education, and Welfare, the European Union's General Data Protection Regulation ("GDPR") in 2016, and in scores of privacy statutes, regulations, and conventions over that 51 year span.

8. **UMass must provide clear guidance as to the use of AI in decision-making, evaluation, and assessment contexts. Use of AI in particular high-stakes or sensitive decision-making contexts should be prohibited.**

   For example, AI tools should never be used to assess the quality of scholarship or student work, to assess merit for financial aid, or to compare candidates for jobs, among other high-impact or sensitive areas. Such uses present insuperable concerns about bias, as well as accuracy, accountability, and transparency / openness. Moreover, even if there is no apparent or obvious effect, there may be less-obvious or second-order bias effects.

   Some clear examples include (but are not limited to):
   - Student related actions and assessments, including admissions, financial aid, instructional activities such as grading coursework, and advising
   - Personnel actions, such as hiring, promotion, tenure, assessment, and discipline
   - Assessments for potential disciplinary or criminal scrutiny

   This guidance also requires a robust educational initiative available to all staff, to understand how and when – and when not – to use AI tools in analysis or other decision making support contexts. This must include the risks relating to bias, as well as to larger concerns relating to accuracy and reliability.

9. **UMass must provide clear guidance on the use of interactive AI or AI-generated works in communications.** For example, the use of AI to write a letter of recommendation or an expression of grief are always inappropriate, because it degrades the human connection that such circumstances demand. Similarly, while the use of AI to conduct interactive conversations may be appropriate in low-level phone trees, it would be inappropriate in situations in which a person needs or expects empathy, human judgment, or connection.

10. **The use of AI in any context – communication, generative, analytical, or otherwise – should be guided by principles of transparency.** The University must develop policy to make this clear. While this principle is clearly drawn from academic norms and therefore intuitive in academic contexts, it is also important in non-academic contexts – such as interactive communications, official communications from the university, and the wide array of business operations the University conducts.

11. **Education in Balancing Risks: The University must develop educational programs to support staff in understanding and using AI appropriately.** Beyond the clear guidelines that we recommended to be set forth above, we also note that **any** use of AI should be carefully assessed, balancing the potential benefits of that use with the potential or known risks. To do so, staff need education about AI – how it works and doesn't work, what the risks and concerns are – in order to appropriately assess whether or not to use it in a given situation.

    The University has a duty to provide the resources to staff to understand how AI tools function, and the risks of bias and other inaccuracies, privacy and data security, and the larger ethical and social concerns that AI raises. AI functions have increasingly been added to other workforce tools – including transcription software, word processing, and search engines, to name a few. All staff – administrative, academic, faculty, graduate employees – are already encountering AI tools. They need the resources to assess the risks, understand the potential benefits, and can therefore appropriately balance the costs and benefits for use in a given situation.

12. Finally, we note that the possible impact of AI poses both a **duty to educate our students** and an **opportunity to bring UMass' values to this area.**

    UMass is a leader among universities in sustainability, social justice, and ethical concerns, and in particular its expertise in critical media / information studies and public interests in technology. We are well-positioned to establish UMass as a leader in assessing these concerns – and, it is our opinion that this is one of the ways we will best serve our students.