# Advancing AI Research at the University of Virginia

Report of the
Provost's Task Force on AI Research

31 March 2024

**University of Virginia Provost's Task Force on AI Research**


*Task Force Co-Chairs*

**David Evans**
*Professor of Computer Science*

**Rafael Alvarado**
*Associate Professor of Data Science*

*Task Force Members*

**Madhur Behl**
*Associate Professor, Computer Science, Systems and Information Engineering*

**Hudson Golino**
*Associate Professor of Psychology*

**Yael Grushka-Cockayne**
*Altec Styslinger Foundation Bicentennial Chair in Business Administration*

**Jingjing Li**
*Andersen Alumni Associate Professor of Commerce*

**Jon Michel**
*Director of Data Science at University of Virginia Health*

**Ricky Patterson**
*Director, Research Data Services & Social, Natural, and Engineering Sciences, UVA Library*


**AI-generated image of the committee hard at work in Shannon Library.**

Any resemblance to actual committee members or a real conference room is purely an AI hallucination.

# Contents

# Advancing AI Research at the University of Virginia

Report of the Provost's Task Force on AI Research
31 March 2024


**Executive Summary**

Advances in machine learning have transformed research across many disciplines, and a large fraction of UVA's faculty are now doing research that involves artificial intelligence and machine learning (AI). UVA has potential to be at the forefront of both core AI research, which seeks to develop the science and technology and ethics underlying AI tools and their use, as well as research that uses AI to advance other areas, including in sciences, humanities, and medicine. Many UVA researchers are among the leaders in these areas, and individual units at UVA have made substantial investments in AI research, primarily through faculty hiring. We lack, however, a dedicated centralized effort to organize, grow, and promote research involving AI, and most UVA researchers have only limited ability and insufficient resources to incorporate AI in their research. Our main recommendations call for creating a Center for AI Research at UVA, and suggest activities of such a center.


## 1 Introduction

Over the past 500 years, traditional automation has made progress to the point where any information processing task that is understood well enough by humans so that the collective efforts of humanity can precisely describe how to do it can be executed quadrillions of times faster than it could be done manually by any human, can be scaled trillions of time larger than humans could do it, and can be executed with near-perfect reliability and remarkably low cost.

What machine learning does is enable automated systems that are *trained* on data rather than *programmed*. With the right learning algorithm and model architecture, machine learning can train a system to do things that humanity collectively does not understand how to do. Machine learning is not a new idea—it goes back to at least the 1950s—but advances in computing power, data availability, and algorithmic breakthroughs and engineering developments over the past decade have dramatically expanded the capabilities of what can be done with machine learning. The term "AI", traditionally an acronym for "artificial intelligence", is used to mean many different things these days and there is no precise and widely accepted definition, but it is now commonly used interchangeably with machine learning, meaning building computer systems by training on data rather than human-engineered programming.[1]

Much of the excitement of the past few years has centered around *generative AI*, familiar to many people through ChatGPT. Traditional machine learning models product just a small, discrete output. For example, an image classifier takes as input an image and outputs a whole number identifying the predicted class of the image (e.g., is it a "cat" or a "dog"?); it may also output a floating point number that indicates the confidence in the prediction. Generative models produce more interesting outputs — an image generator produces a high

[1]The other notion of *artificial intelligence* as building mathematical models to understand and replicate human thinking has an even longer history, going back at least to George Boole's work in the 1850s. Although there is still and active and vibrant research community focused on this goal, most work called AI today is focused on building systems that solve practical problems, with only loose connections to how learning and reasoning work in biological systems.

resolution image, a language model produces English text, a protein structure model predicts the locations of atoms in a protein.[2]

Another cause for recent exuberance is the increasing generality of machine learning models. This is sometimes referred to as *artificial general intelligence* (AGI), and questions about whether or not current models exhibit "intelligence" or if current approaches are on a path towards achieving truly general intelligence engender intense philosophical debates. We avoid those issues here, but note that recent models seem to be capable of performing a large number of different tasks which they were not explicitly trained to do, and being capable of quickly being tuned or trained to perform new tasks.

**Scope of the Task Force.** Although we think having a precise notion of AI and understanding of recent developments in generative AI is important for understanding the opportunities in AI and AI-related research, we did not limit our scope to a particular definition of AI or narrow focus on recent generative AI developments. In our survey and discussions with people around the University and outside, we considered AI broadly and were open to anything anyone involves considered relevant to AI. We use the term *AI-based research* throughout the report to include both types of core AI research and the use of AI in service of all research.

## 1.1 Task Force Development and Charge

The task force was initiated in October 2023, and the two co-chairs (Rafael Alvarado and David Evans) were asked and agreed to lead the task force in early October. The initial questions posed at this time were:

- *What does the Provost and VPR need to know to support the research community regarding AI?*
- *Are there investments we need to make?*
- *Is research computing appropriately aligned with AI?*
- *Is there a subfield of AI that we could stake out where we could be leaders?*
- *Are there fellowships that are needed?*
- *Do we need an AI core?*
- *What resources should be made available?*
- *Are there other key issues?*

The co-chairs were asked to suggest a list of members for the committee, and submitted a list of ten potential committee members, covering a broad perspective of the university and range of expertise and career levels, to the VPR and Provost representative in early November. The VPR/Provost selected four candidates from the co-chairs' list, and suggested two other members (one of the committee members was suggested later in the process, not in the initial list). The co-chairs invited all of the suggested members to the committee, and everyone promptly agreed to join!

This resulted in the final Task Force membership listed below:

- Rafael Alvarado (co-chair, Data Science)
- Madhur Behl (Computer Science, Systems and Information Engineering)
- David Evans (co-chair, Computer Science)

[2] This distinction, like most in this field, is fuzzier than it seems. Most language models just predict the next word, one at a time. Each next word prediction is more like what an image classifier does—outputting one word from a dictionary of possible words.

- Hudson Golino (Psychology)
- Yael Grushka-Cockayne (Darden)
- Jingjing Li (McIntire)
- Jon Michel (Health)
- Ricky Patterson (Library)

In addition to these members, Fred Epstein, Interim Vice President for Research and Mac Wade Professor of Biomedical Engineering, and Cheryl Wagner, Chief of Staff to the Vice President for Research, contributed throughout the process, with at least one of them joining most committee meetings and providing insights and logistical support throughout the study.

The committee charge was refined over the following several weeks, resulting in the final charge questions below. The committee had two meetings before the new year, and began meeting regularly the first week of January, holding meetings each week since then to coordinate progress.

The task force was charged with responding to the following questions:

1. *Hardware, software, data, human and organizational resources – Is the University providing adequate resources for UVA leadership/competitiveness in AI research across grounds?*
2. *Should the University be providing training for all UVA researchers in AI ethics, fairness, bias and related areas? Make recommendations about the type and modes of training.*
3. *Do UVA policies on research and data need to be updated due to AI?*
4. *Use of AI to facilitate the UVA research enterprise – should VPR research development start to develop AI models to be able to mine our own data (papers, submitted grants, IRB protocols, animal protocols, IP, etc) in order to optimize research development at UVA? Others are already doing this with impressive results, and we may fall behind without such an effort.*
5. *What does UVA need to do to be considered competitive for programs like NSF AI Research Institutes or the USAISI?*
6. *Anything else the committee brings to light.*

### 1.2 Committee Activities

To address the charge questions, the task force has undertaken a range of activities, beginning in earnest in January 2024. The main information gathering activities conducted include:

- Conducted a survey of a sample of UVA community members to learn about current AI research, plans for upcoming research, and other topics related to the charge questions. The survey was prepared in December 2023 and sent out to a selection of department chairs in January 2024 (we were not able to survey the full community due to University policies and the tight timeline for this task force). We received over 230 responses. Section 2.1 explains the survey methods and analyzes the survey responses using both quantitative and qualitative methods.
- Based on the survey results, we invited all the respondents who indicated they would like to follow-up with more discussion and provided an email address to join an open forum discussion. Two forums were conducted in person, and two on zoom, from 29 January through 5 February. Forums were attended by a broad cross-section of faculty across the university, and the discussions were engaging

and enlightening. We discuss the forums further in Section 2.2.

- To understand current resources and plans for Research Computing, we met with Joshua Baller and received further information from the VPR office.
- We conducted a bibliometric analysis of publications at UVA and peer institutions, using data from OpenAlex, to understand publication trends and how UVA compares to other institutions (Section 2.3).
- Investigated AI initiatives and resources at other universities, primarily through what is available from their public websites and press releases (Section 2.4).
- To respond to charge question #5, we have had discussions with faculty involved in NSF AI Institute proposals (and several task force members have direct experience with both successful and unsuccessful large scale proposals including NSF AI Institutes). We also analyzed results from previous AI Institute competitions (Section 2.5).
- To understand issues relevant to the ethics and policy charge questions, we surveyed activities at other universities and conducted a number of interviews (Section 2.6).
- To understand research enterprise operations, we met with suggested representatives of the VPR office and have collected information on data sources used by the VPR office. Kim Mayer also provided us with notes from her conversation with Rob Rutenbar, Senior Vice Chancellor for Research at the University of Pittsburgh (Section 2.7).

Based on the information collected and analyzed, we discussed and debated our findings and recommendations. We had approximately weekly zoom meetings, and an extended in-person recommendations workshop to develop our recommendations. We developed the report using Overleaf, a collaborative writing tool, and all task force members reviewed and agreed to the recommendations contained in this report.

## 2 Landscape

To inform our recommendations, we started by gathering information on the current landscape of AI-based research at the University, at peer and aspiring peer institutions, and broadly across government and industry. Section 2.1 summarizes what we learned about research and planned research at UVA from a community survey. Section 2.2 describes the meetings we held and what we learned from them. Section 2.3 summarizes our analysis of research publications. Section 2.4 describes initiatives underway at other universities, and Section 2.5 looks at the NSF AI Institutes. Section 2.6 covers AI ethics and Section 2.7 provides background in use of AI to support the research enterprise.

### 2.1 Survey

In January 2024 we disseminated a survey across various university units to discover facts and opinions about the use of AI in research among faculty. Because of university policies that restrict the broadcasting of email messages to faculty and the need to move quickly we were not able to send a survey to the entire community. Instead we sent messages to a list of department chairs and asked them to forward the survey request to their faculty.

We received 233 responses to the survey (this corresponds to about 15% of the approximately 1600 faculty at UVA. (We cannot estimate the actual response rate relative to requests, since the number of people actually receiving a request to submit the survey depends on how chairs receiving our request distributed it). Although our sample is not a random sample of the UVA population due to our inability to send a survey request to the full community, it is large and broad enough to draw some conclusions based on the data we received. Of the survey respondents, 188 (81%) completed the survey. The survey was anonymous, but at the end respondents were invited to provide contact information if they wanted to follow-up with the task force. Of the 188 respondents completing the survey, 56 indicated wanting to follow-up and provided valid contact information (these were invited to the Faculty Forums, see Section 2.2).

In this section, we summarize what we've learned about research activities and plans at UVA from the survey; we defer the survey results focused on ethics to Section 2.6.1.

**Diversity of disciplines and myriad use cases.** The survey results show a remarkable range of disciplines interested in AI. Respondent affiliations include the humanities, arts, social sciences, natural sciences, engineering, and many of the professional schools. Table 1 summarizes the respondents. Although the majority of the responses are from people in STEM fields, approximately a third of respondents are from members of non-STEM fields. The diversity of disciplines is matched by the diversity of approaches to and applications of AI found in the survey responses. Use cases range from using AI to advance primary research in the physical sciences and engineering as well as in medicine and the humanities to artistic uses of AI tools. Many faculty are also involve in critically evaluating AI developments from ethical, philosophical, and social perspectives. Table 2.1 and Appendix B.3 provide some more depth on the areas where faculty at UVA are doing AI-based research.

**Core, applied, and critical research in AI.** There is a broad division between core AI researchers, who aim to understand and advance AI tools and methods; users of AI, who use AI to advance research in other disciplines; and those whose interest in AI is mainly critical and evaluative.

Table 1: Response rates by school and unit. (Note that providing an affiliation was optional in the survey, and respondents provided their affiliation in free text so there is some ambiguity as to actual affiliations.)

| School | Count | Units Mentioned |
|---|---|---|
| Arts & Sciences | 75 | **Humanities** (30): English (13), Music (4), Writing and Rhetoric (2), African American and African Studies (2), German (2), French (2), History (1), Creative Writing (1), East Asian Languages, Literatures and Cultures (1), Women, Gender, and Sexuality (1), Art and Archaeology (1)<br>**Natural Sciences** (26): Mathematics (10), Physics (8), Biology (4), Astronomy (3), Statistics (1)<br>**Social Sciences** (12): Psychology (6), Sociology (3), Economics (3)<br>Unknown (7) |
| Engineering & Applied Science | 51 | Biomedical Engineering (12), Computer Science (10), Engineering and Society (8), Materials Science and Engineering (4), Mechanical and Aerospace Engineering (4), Civil and Environmental Engineering (2), Systems and Information Engineering (2), Chemical Engineering (2), Electrical and Computer Engineering (2)<br>Unknown (5) |
| Medicine | 24 | Public Health Sciences (6), Cell Biology (5), Microbiology, Immunology and Cancer Biology (3), Center for Public Health Genomics (1), Pediatrics (1), Neurology (1), Carter Immunology Center (1)<br>Unknown (6) |
| Business | 8 | Quantitative Analysis (1), Strategy, Ethics, and Entrepreneurship (1)<br>Unknown (6) |
| Education & Human Development | 7 | Curriculum, Instruction, and Special Education (2), General (2), Kinesiology (2); Unknown (1) |
| Data Science | 5 | |
| Commerce | 4 | IT and Innovation (1); Unknown (3) |
| Law | 1 | |
| Nursing | 1 | |
| Other | 5 | College at Wise (2), Biocomplexity Institute (1), Office of the VPR (1), University Library (1) |
| None Provided | 52 | |

**Division of AI use.** A majority of respondents (71%) expect to maintain, increase, or centralize their use of AI in the future, although non-users of AI (41%) and those who consider AI of minor importance (51%) were well represented. The data suggest a division among those who have either applied AI to their research or whose research focuses on it, and those who aspire to apply AI to their areas. The former group have a high number of publications related to AI and come from the sciences, engineering, and the professional schools. The latter come from a wider variety of fields and have expressed a broad range of research interests in their descriptions of research goals and possible projects. The distribution of output and provenance among the groups appears to follow a power law distribution, raising concerns that a rich-get-richer effect may follow from allocating resources to areas that have already demonstrated success, where allocating resource to underdeveloped areas may have more impact. Recommendation 9 aims to provide more opportunities for researchers with ideas but without previous experience or knowledge of AI to have opportunities to learn and develop research programs that leverage AI in novel ways.

**Great expectations but little comfort.** Respondents have high expectations for AI's potential to advance research, even though many who perceive its potential do not count themselves as comfortable with AI. There is a very high aggregate correlation (Pearson correlation coefficient, $r = .92$) between the sample population of those who have a high comfort level with AI and the sample population of those who express interest in its potential to enhance their research. However, this is not reflected at the individual level—comfort level with AI and its perceived potential have a moderate positive association ($r = .55$). This implies that not all of those who find potential in AI are comfortable with it, and not all who are comfortable with it think it has potential for their research. Appendix B.1 provides more details on this analysis.

**Publications.** The survey asked respondents to describe their AI-related publications, and out of the 233 responses there were 113 respondents who answered the question "How many papers have you published that involve or use Artificial Intelligence?". Of these, 68 (60% of the 113 responses; since respondents could leave this question blank, it is likely that most of the respondents who did not answer this question did not have any AI-related publications) indicated no AI-related publications yet, 45 (40% of the question responses, or 19% of the overall survey responses) indicated at least one publication and 11 respondents indicated more than 10 publications.

Table 2 shows the number of publications in each topic, selected from a prescribed list in the survey, reported by respondents. Machine learning and neural networks (a particular approach to machine learning that developed in the 1960s, but has had great success over the past decade) account for over half (53%) of the publication output, which is roughly seven times more than the rest of the topics combined and 2 to 15 times more than any individual topic. This suggests that the current output of research in AI is heavily weighted toward a relatively small group of machine learning researchers. We also asked respondents to provide free-text responses to describe their AI publications and to provide a description of a selected publication—the responses on these questions are summarized in Appendices B.2.2 and B.2.3. The free text responses exhibit a dazzling diversity of research consistent with the diversity of disciplines represented by the respondents. Topics range from quantum computing in physics to explainable AI to AI in popular culture and the arts—almost any research topic in any discipline appears to be ripe for applications of AI in one form or another.

The bibliometric analysis (Section 2.3) provides a more comprehensive and objective measure of UVA's publication activity in AI, but is mostly consistent with the survey responses.

Table 2: Number of AI-related publications reported by survey respondents.

| Topic | Publications |
|---|---|
| machine learning | 60 |
| neural nets | 31 |
| healthcare | 28 |
| computer vision | 21 |
| ethics | 20 |
| NLP | 17 |
| robotics | 14 |
| security | 13 |
| education | 13 |
| other | 13 |
| finance | 6 |
| Total | 170 |

### 2.2 Forums

Based on the survey responses, four discussion forums were setup to have more in-depth discussion. We sent invitations to the forums to the 56 respondents to the survey (out of 188 who completed the survey) who indicated that they wanted to follow-up with the task force and provided contact information in their submission. We sent email to all these respondents inviting them to join any of the scheduled forums or to reach out to us directly if they were not able to make it to a forum. The forums were scheduled between 29 January and 5 February 2024, with two of them held in person (one in Clark Hall Library and one at Darden) and two held on zoom. Invitees were free to join whichever forum was convenient for them.

Each forum lasted an hour and the discussion was facilitated by a task force member, with at least one and typically two or three additional task force members joining. Although the meetings were mostly unstructured, the following questions were used to frame the conversation:

- *Tell us how you use AI in your research*
- *What is exciting to you about how AI is changing your field?*
- *What have you seen UVA do well? Your school/department?*
- *What help do you need from UVA?*
- *What do you think we could do with industry?*
- *What would motivate you to collaborate with others on grounds?*
- *Any equity concerns? Who is drawn to use the tools?*

A total of seventeen faculty (excluding committee members) participated in the forums, with representative from across various schools and departments at UVA. The overall sentiment was of excitement, curiosity, and embracing AI and GenAI, which perhaps can be explained by noting that all those who joined did so voluntarily and out of genuine interest. Some faculty, while eager to learn, also reported that they felt overwhelmed by AI and by the growing popularity and usage of GenAI around them.

Faculty attending the forums described how AI was changing their disciplines. Across domains, transparency and explainability, ethics, and environmental costs were of concern and interest. Faculty also described how research topics related to AI were being fostered by dedicated hiring lines (e.g., a Chair in AI and Creative Writing based on donor request), but at the same time there was some concern that the source of such talent what still limited and uncertain.

The examples of work related to AI represented a broad and diverse topics, including using AI tools to analyze political rhetoric on-line as well as studying the impact of generative AI on political disinformation campaigns; in astronomy, developing large models of the universe from massive data source and using these to improve understanding of dark matter; in computer engineering, using AI to design hardware as well as designing computing systems specifically for AI tasks; in literature, the potential impact of AI on monolingualism; in materials science and engineering, using AI to design advanced materials including automated labs; in medicine, using ML methods for drug discovery and cancer research; in economics, using data-derived models to understand problems in developing countries; in public health, using AI models to detect bacteria in water samples; in law, using natural language processing techniques to analyze supreme court of India judgments; using AI tools to develop nutrition guidance; evaluating psychological aspects of AI systems and testing for implicit biases; in environmental science, developing AI methods for earthquake prediction; and uses of AI in education including helping both teachers and students. Given the separate task force focused on AI in Education, we tried to focus these conversations on research but several faculty expressed interest in experimenting with and using generative AI in the classroom but also raised concerns its shortcoming and risks including inaccurate information and sourcing.

When asked about what UVA is doing well with regards to AI research, there were positive reflections on low-cost efforts hat bring interdisciplinary groups together (an example that came up was the SEAS Research Interest Groups which enabled small groups of faculty to easily obtain funding to support meetings focused on a shared interest).

Interesting discussions emerged on the question of whether changes are required to UVA's Promotion and Tenure criteria in order to facilitate and incentivize more research in AI. Several participants observed that the expectations for faculty in their discipline discourage participation in interdisciplinary work and fail to recognize time-consuming but very valuable efforts in building tools and data sets that are essential for much work in AI but do not result in the kinds of products that are traditionally recognized in most tenure reviews. Participants advocated for changes in UVA's expectations for pre-tenure faculty that rewards for contributing to large scale collaborations, and well as incentives to develop a researcher's capabilities in new directions that may involve periods of lower productivity while new skills such as those needed for advancing AI-based research. Recommendation 9 and Recommendation 18 are partly motivated by these discussions.

We asked faculty explicitly about what resources would be desired to enable them to more effectively use AI in their research. Faculty mentioned interest in university-wide license for generative AI tools (ChatGPT was specifically mentioned), large storage capabilities with high throughput and the ability to support petabyte-scale datasets, substantial GPU compute resources, facilitate GPU computing with industry partners, support for visiting researchers from industry and other universities, support for post-doctoral scholars to kick-start AI research programs, high-profile hiring of leading experts in the use of AI, and agility in order to move into areas more quickly than federal funding cycles. Faculty also suggested that smaller grants for forming pan-university project teams could be useful, especially in the humanities and social sciences. Other asks included the need to offer competitive salaries to recruit talent in AI, as well as the need for even more collaborations across grounds, institutionalizing joint appointments, supporting curriculum development, and creating physical spaces for creative work and meetings.

Issues regarding ethical use of AI were also raised during the forums. Concerns were raised about AI's interpretability, the accuracy of AI-generated content, and the ethical implications of AI in the context of education and research. Some participants advocated for a university-wide license for AI tools to ensure equitable access. The importance of explainability was a cross-cutting theme across disciplines, evoking lively discussions about the nature of the problem.

Forum attendees repeatedly mentioned that there could be improvements made for how innovations were communicated across grounds and for streamlining and ensuring all are aware of developments and opportunities (as one concrete example, many researchers were not aware of the consultation and collaborations service available through the Data Analytics Center, but those who knew of them spoke highly of how useful they were). Although there was not widespread agreement on how to most effectively communicate with the faculty and multiple channels will be needed, the main mechanisms discussed were improving websites to better reflect the needs of users instead of the organizational structure of the university, and sending out periodic and carefully curated newsletters about research activity and opportunities. Faculty were interested in more workshops on AI, for students and for faculty.

With regards to collaboration with industry, participants felt that industry is leading the AI revolution (perhaps recklessly), and that certain types of work can only be done in industry or in close collaboration with industry. There were suggestions that perhaps UVA faculty could spend time in industry, to allow for funding for research. Perhaps embracing a sabbatical model, or a partnerships with PhD programs.

### 2.3 Bibliometric Analysis

To understand trends in research, and how UVA compares to other institutions, we analyzed bioliometric data on publications involving AI and related areas. For the comparison, we focused on the universities identified by the University's Strategic Research Initiative as "peers and aspiring peer institutions for use in benchmarking UVA". The tweleve instutions are: Duke University, Emory University, Northwestern University, Ohio State University (OSU), Rutgers University, the University of Alabama at Birmingham (UAB), the University of Michigan, University of North Carolina at Chapel Hill, the University of Pittsburgh, the University of Washington, Vanderbilt University, and Virginia Tech.

Publication records were extracted using OpenAlex (https://openalex.org), an open catalog of publications that indexes over 250 million works from 250,000 sources. OpenAlex links these works to 90 million disambiguated authors and 100,000 institutions, and provides information on publications including citation counts. The publication records were extracted using as search indexes the following concepts: "artificial intelligence", "machine learning", "computer vision", "neural networks", and "natural language processing". This resulted in 294,087 publication records for analysis.

We used machine learning techniques including large language models to analyze the publication data (Appendix C provides details), identifying a set of topics and assigning for each paper a score the reflects how relevant the paper is to that topic. Figure 1 shows the number of papers published in "Artificial Intelligence" and "Machine Learning" from 1980–2023 (these are distinguished as topics in the paper analysis, but, as discussed in the Introduction (Section 1), are used interchangeably in many contexts, including this report).

The two main observations from these graphs are that research in AI has increased spectacularly since about 2015, and that UVA appears to be falling behind our peers. As measured by publication count (which of

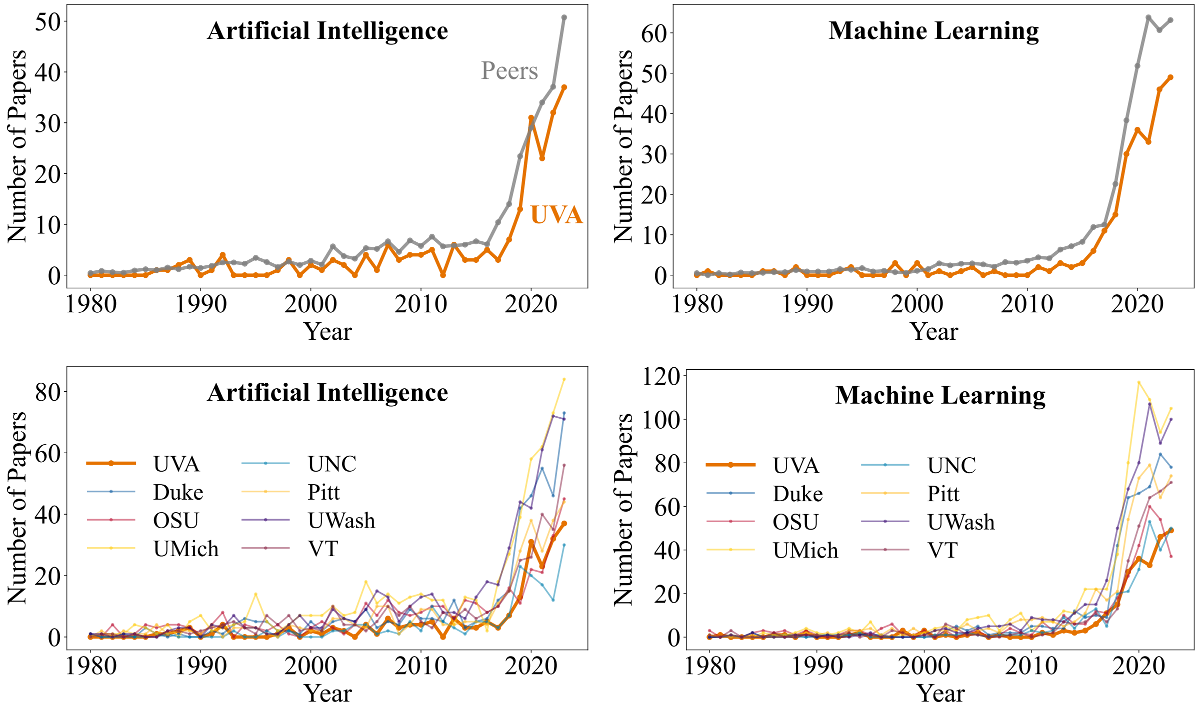


Figure 1: Number of Papers in the OpenAlex data in AI and ML published by UVA and peer universities (based on high confidence topic score), 1980–2023. The grey line in the top graph is the average for all of the peer universities (excluding UVA); the bottom graphs show a selected subset of the peer universities. The dip in 2023 reflects the data for 2023 being incomplete in our dataset. A paper is counted for all institutions where at least one author of the paper is affiliated with the institution, and is counted in all topics for which its topic score indicates high relevance.

course is a despicable way to measure research value, but is a representative and objective measure), UVA was roughly in the middle of this peer group from 1980–2012, but has fallen behind, and now ranks near the bottom. For the data available for 2023, UVA ranks twelfth (out of 13, ahead of only University of North Carolina among the peers) in Artificial Intelligence publications and ninth (out of 13, ahead of University of Alabama at Birmingham, Ohio State University, Rutgers University, and Vanderbilt University) in Machine Learning.

We also analyzed how research efforts are distributed at different institutions, and summarize those results in Figure 2. The figure shows the mean score for each of the six topics shown for each paper in the data set for UVA and the average across the peer institutions. UVA developed an early focus in bioinformatics in the late 1980s (largely due to Bill Pearson's pioneering work on genetic data analysis), but otherwise has largely followed similar trends as the peer institutions.

The data for these figures covers a broad range of publications across all areas, selected based on the AI-relevant concepts as provided in the OpenAlex data. We also wanted to understand trends in top-tier venues for publishing AI research, so analyzed available data on publications in the conferences that are generally regarded as the most important places for publishing AI research: the *Conference on Neural Information Processing Systems* (NeurIPS, first held in 1987), the *International Conference on Machine Learning* (ICML,

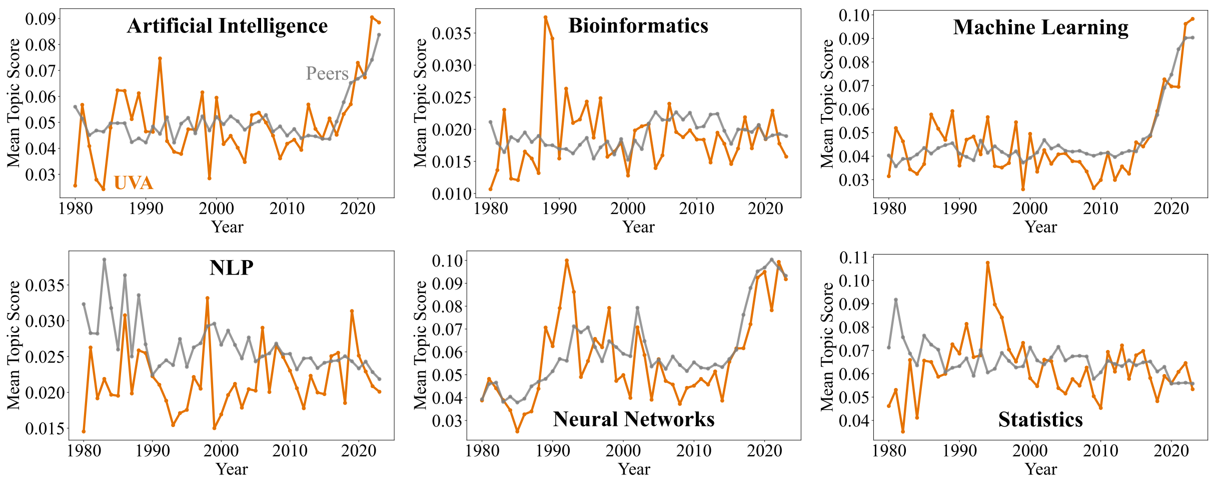


Figure 2: Mean Topic Scores for several AI-related topics, comparison between UVA and peer universities. Note that the vertical axes are different scales for the graphs, reflecting the different levels of activity (as well as how specialized the topics are) in each area.

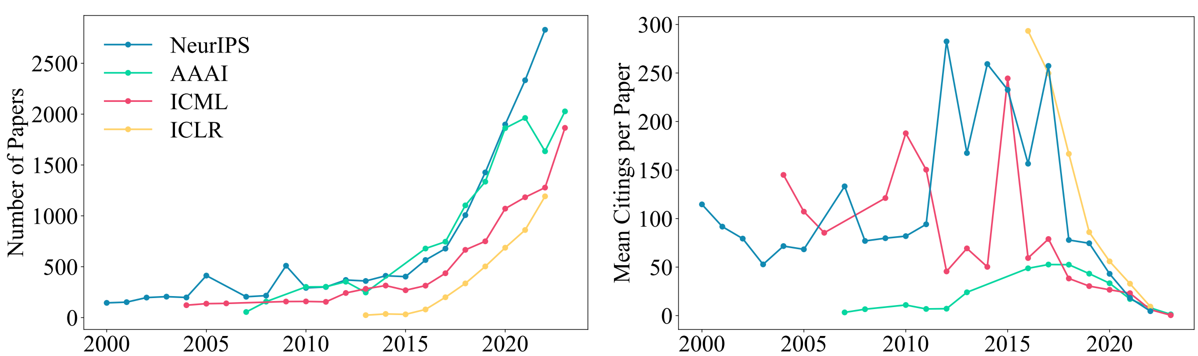


Figure 3: Growth of AI Conferences. The conferences shown are widely regarded as the top venues for publishing AI research: *Conference on Neural Information Processing Systems* (NeurIPS), *AAAI Conference on Artificial Intelligence* (AAAI), *International Conference on Machine Learning* (ICML), *International Conference on Learning Representations* (ICLR, first held in 2013). The left graph shows the total number of papers in that year's edition of the conference. The right graph shows the mean number of citings for papers in the conference (data for ICLR is not shown for 2015 and earlier, since the number of papers was low the results are highly distorted by a few very highly cited papers—the average for ICLR 2015 is over 2500 citings per paper, because of two papers with over 20,000 citings each). Citings of recent conferences are lower because there has been limited time for those papers to be cited (and additional lag because of the time it takes for papers to be recorded in Scopus dataset).

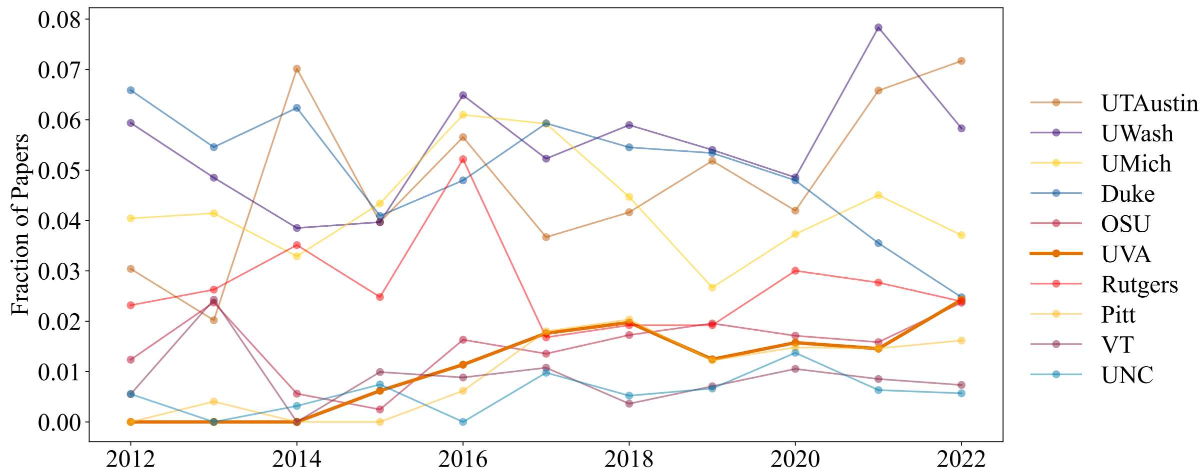


Figure 4: Publishing in Top-Tier AI Conferences. The fraction of the total papers across major AI conferences with authors with affiliations with the comparison universities.

first held in 1980), the *International Conference on Learning Representations* (ICLR, first held in 2013), and the *AAAI Conference on Artificial Intelligence* (AAAI, first held in 1980). We collected data from these conferences using Scopus, a commercial research publication service available through the University library's subscription.

Figure 3 shows the remarkable growth in the sizes of the top-tier AI conferences, reflecting the overall growth in the field since 2012. The most recent version of NeurIPS (2023) received 12,345 submissions of which 3218 (26%) papers were accepted. (The 2023 conferences are not included in Figure 3, since Scopus data does not include all of the 2023 conferences as of March 2024.) This acceptance rate is typical for these conferences, and papers are selected using a rigorous peer review process.

Figure 4 shows the fraction of papers in the top AI conferences that include at least one author from the designated institution. UVA's publications in these venues show significant relative growth between 2014 and 2018, and falls within the middle of the group of comparison universities (UT Austin was added to the comparison here, but is not included in the peer universities list).

Although the results are mostly consistent with the analysis of the OpenAlex data using the topic model, this data only includes the most visible AI-focused conferences, where the OpenAlex data includes a broad set of publications across many venues, including both less prestigious AI conferences and venues that are not focused on AI research. UVA's better showing in this data could be interpreted as reflecting our relatively strong research in core AI, but comparative weakness in applying AI methods to other research areas. It is also noteworthy that of all the comparison institutions, UVA has made the most progress since 2014 (when we ranked as the bottom institution among this group). Some may also appreciate that whereas Duke was ranked first in 2012 and UVA last, by 2022 the results for UVA and Duke are effectively indistinguishable.

## 2.4 AI Initiatives at Other Institutions

Over the past decade, and accelerating in the past few years, many universities have made substantial investments in AI. Some notable examples include:

- **Johns Hopkins University** (JHU) launched a new *Data Science an AI Institute*[3] in August 2023 with a plan to hire 80 tenure-track faculty in AI within the next five years. The plan includes 30 new "Bloomberg Distinguished Professors" with an interdisciplinary focus, and a projected 750 PhD students, as well as 30 scientific staff including research engineers. The institute is aligned with one of JHU's ten strategic goals[4]: "Create the leading academic hub for data science and artificial intelligence to drive research and teaching in every corner of the university and magnify our impact in every corner of the world.".
- **Princeton University** announced the *Princeton Language and Intelligence Initiative* in September 2023.[5] The initiative was announced as part of a New Jersey state initiative to become an "AI Innovation Hub", and was described by Princeton's Provost as reflecting the institutions three strengths in foundational AI research, a commitment to interdisciplinary research, and a commitment to service.[6] The initiative emphasized investments beyond typical faculty hiring, with support for research software engineers and research scientists, and a substantial investment in computing resources (300 Nvidia H100 GPUs, with a current market value of approximately $10M).
- The **University of Texas at Austin** declared 2024 to be the "Year of AI".[7] One of several major research initiatives that are part of this is the *enter for Generative AI*, which includes a GPU cluster with 600 Nvidia H100 GPUs. In announcing the center, UT President Jay Hartzell said, "Artificial intelligence is fundamentally changing our world, and this investment comes at the right time to help UT shape the future through our teaching and research... World-class computing power combined with our breadth of AI research expertise will uniquely position UT to speed advances in health care, drug development, materials and other industries that could have a profound impact on people and society."[8]

Although these universities are not on the list of peers identified for benchmarking by the Strategic Research Initiative, and some may consider them in a different league than UVA, we point out that UVA's annual research expenditures exceed Princeton's[9], our endowment is similar to Johns Hopkins' endowment[10], and UVA was founded 64 years before UT Austin and is ranked well above them in most university rankings. Although these universities have advantages over UVA in some respects, we see no reason for UVA's ambitious to be so diminished to consider similar prominence to be beyond our reach.

#### 2.4.1 Other Universities in Virginia

We also investigated AI research and strategy initiatives across the other four R1 universities in Virginia: Virginia Tech, George Mason University (GMU), Old Dominion University (ODU), and Virginia Commonwealth University (VCU). Our goal was to understand the investments each university is making, and the

[3] https://ai.jhu.edu/, https://hub.jhu.edu/2023/08/03/johns-hopkins-data-science-artificial-intelligence-institute/, personal emails from faculty at JHU.

[4] https://president.jhu.edu/ten-for-one/

[5] https://pli.princeton.edu/

[6] https://www.princeton.edu/news/2023/12/20/governor-murphy-and-princeton-announce-plans-establish-artificial-intelligence-hub

[7] https://news.utexas.edu/2024/01/23/ut-designates-2024-the-year-of-ai/, https://yearofai.utexas.edu/

[8] https://ml.utexas.edu/center-for-generative-ai, https://news.utexas.edu/2024/01/25/new-texas-center-will-create-generative-ai-computing-cluster-among-largest-of-its-kind/

[9] According to the latest data we could find (https://ncsesdata.nsf.gov/profiles/site?method=rankingBySource&ds=herd, which is for 2022), UVA ranked 48th with $662M, and Princeton was 63rd with $461M.

[10] UVA's was $13.6B in June 2023 according to https://uvimco.org/annual-report/annual-report-2023/; the most recent information we could find on JHU's listed it as $10.5B (https://universitybusiness.com/the-top-20-university-endowments-of-2023/), which lists UVA's as $9.8B.

strategies they are following regarding focus areas, research infrastructure and investments, partnerships, and community impact. We did not have access to any internal information from these universities, so our analysis is based entirely on information collected from the universities' public websites and related resources.

**Virginia Tech.** Virginia Tech has several major strategic initiatives focused on AI-based research:

- **The Artificial Intelligence Frontier:** Virginia Tech prioritizes AI as a key area of research alongside health, security, and quantum technologies. The university is promoting interdisciplinary collaborations and partnerships in AI. Their *Artificial Intelligence Frontier*[11] is a university wide research initiative, and involves interdisciplinary collaboration across fields such as data science, systems engineering, and neuroscience. The initiative focuses on enhancing human-technology partnerships, with significant projects in autonomous systems, machine learning, and smart cities.
- **Sanghani Center for AI and Data Analytics:** (https://sanghani.cs.vt.edu/) Initially established as the Discovery Analytics Center in 2011, the Sanghani Center focuses on research in several AI areas, covering topics like natural language processing, explainable AI, adversarial AI, and network analysis, and application areas such as public health and national security. The center has over 20 faculty and a large number of graduate students across its two locations in DC and Blacksburg.
- **Amazon–Virginia Tech Initiative:** This partnership, initiated in 2022[12], aims to improve research in efficient and robust machine learning through fellowships and faculty-led research projects. It emphasizes creating machine learning models that are efficient and resilient to errors and adversarial attacks. The initiative includes public symposia, workshops, and events to foster collaboration between Virginia Tech and Amazon.[13]

Virginia Tech has a comprehensive approach, integrating AI across various strategic areas, an established AI research center, and boasting significant partnerships, like the one with Amazon, which position it strongly in terms of innovation and industry collaboration. Virginia Tech clearly identifies AI as a research priority through its dedicated Artificial Intelligence Frontier initiative and web page, underlining the university's commitment and making its focus readily apparent.

**George Mason University.** GMU launched a major center last year that involves both research and education in AI:

- **Mason Autonomy and Robotics Center (MARC):** (https://marc.gmu.edu/) Scheduled to launch a graduate certificate in Responsible Artificial Intelligence in fall 2024, MARC focuses on the risks and benefits of AI and robotics. The center, which opened a new facility in fall 2023, has over 50 faculty members as well as a dedicated Director (Dr. Missy Cummings, UVA PhD alumnae) and Associate Director. It engages in interdisciplinary research and provides education on AI and robotics through GMU's degree programs. The center's goal is to understand and mitigate the ethical and social implications of AI-enabled autonomous systems.

GMU has an AI Strategies Team, funded by a three-year $1.4 million grant, to study the economic and cultural determinants for global AI infrastructures, and describe their implications for national and international security (https://www.aistrategies.gmu.edu/).

[11] https://www.research.vt.edu/initiatives/research-frontiers/artificial-intelligence.html
[12] https://www.amazon.science/academic-engagements/amazon-and-virginia-tech-launch-ai-and-ml-research-initiative
[13] https://sanghani.cs.vt.edu/amazon-virginia-tech-initiative/

**Virginia Commonwealth University.** VCU does not appear to have any university-level initiatives focused on AI research,[14] but does have several school-level efforts and education-focused initiatives, including a *Vertically Integrated Project* on AI Governance[15], a new minor in Artificial Intelligence[16], and an AI Futures lab focused on understanding power relations in AI systems and developing ethical AI.[17]

**Old Dominion University.** ODU has strategic initiatives in Cybersecurity and Spaceflight & Autonomous Systems, but not appear to have any AI-focused research initiatives.[18] It does have a broad effort on integrating AI Content Generation Tools into teaching, learning, and research.[19]

## 2.5 NSF AI Institutes

In 2020, the National Science Foundation initiated a program to fund large-scale AI Institutes, and as of March 2024 there have been 25 AI institutes funded, each at roughly $20 million. Charge question #5 asked us to specifically address how to make UVA more competitive for NSF AI Institutes.

Table 3 shows the universities that lead AI Institutes as well as those who have received subawards. UVA has not received an AI Institute as a Primary (lead) institution, but is a subawardee for two AI Institutes—the AgAID Institute for Agricultural AI for Transforming Workforce and Decision Support (let by Washington State University), and the ACTION Institute for Agent-based Cyber Threat Intelligence and Operation (led by UC Santa Barbara). UVA also has had success with other center-scale NSF funding related to AI, including most notably as lead for the Expeditions Award on *Global Pervasive Computational Epidemiology* (which includes a substantial AI component).[20] Although funded at a lower level, expeditions are more competitive and more visible in most ways than AI Institutes, with only three awarded in 2020 (and two in 2022), compared to 25 AI Institutes awarded since 2020. UVA is also a subawardee in the first major NSF-funded center on trustworthy machine learning (the SaTC Frontier Center on Trustworthy Machine Learning, initiated in 2018)[21]. It is notable that UVA's most notable success leading a major NSF project, as well as one of our two subawards for an AI Institute, both were done through the Biocomplexity Institute, demonstrating the value of research-focused interdisciplinary institutes.

## 2.6 Ethics

Although most of the individual areas of concern raised by AI may not be new, their combination and scale is. The new breed of generative AI techniques rely heavily on very large sets of data and have large model capacities that are trained in ways that are prone to memorizing aspects of the training data. For example, a language model may generate verbatim text from its training data in response to a prompt. This text may contain personally sensitive information or be under copyright. This reliance on data introduces ethical problems into the AI space inherited from problems already addressed by many organizations already focusing on data and society.

[14] https://research.vcu.edu/resources/institutes-and-centers/
[15] https://vip.vcu.edu/teams/
[16] *VCU College of Engineering to offer six new minor programs*, 19 January 2024.
[17] https://humanitiescenter.vcu.edu/labs/ai-futures-lab/
[18] https://ww1.odu.edu/research
[19] https://www.odu.edu/facultydevelopment/ai-content-generation-tools-teaching-learning-and-research
[20] https://www.nsf.gov/news/special_reports/announcements/032420.jsp, *Biocomplexity Institute Wins $10M Grant to Thwart Future Pandemics*, UVAToday, 25 March 2020.
[21] https://www.nsf.gov/news/news_summ.jsp?cntn_id=296933

| Institution | Primary | Sub |
|---|---|---|
| University of Illinois (UIUC) | 3 | 5 |
| Georgia Institute of Technology | 3 | 2 |
| Ohio State University | 2 | 0 |
| University of Washington | 1 | 7 |
| Carnegie Mellon University | 1 | 4 |
| MIT | 1 | 2 |
| North Carolina State University | 1 | 2 |
| University of Texas Austin | 1 | 2 |
| Columbia University | 1 | 1 |
| Iowa State University | 1 | 1 |
| UC Davis | 1 | 1 |
| UC San Diego | 1 | 1 |
| Duke University | 1 | 0 |
| UC Santa Barbara | 1 | 0 |
| University at Buffalo | 1 | 0 |
| University of Colorado | 1 | 0 |
| University of Maryland | 1 | 0 |
| University of Minnesota | 1 | 0 |
| University of Oklahoma | 1 | 0 |
| Washington State University | 1 | 0 |

| Institution | Subawards |
|---|---|
| Harvard University | 5 |
| Cornell University | 4 |
| UC Berkeley | 4 |
| University of Wisconsin | 4 |
| Colorado State University | 3 |
| Pennsylvania State University | 3 |
| Purdue University | 3 |
| University of Michigan | 3 |
| Yale University | 3 |
| Indiana University | 2 |
| Northeastern University | 2 |
| Oregon State University | 2 |
| Texas A&M | 2 |
| Tuskegee University | 2 |
| University of Chicago | 2 |
| University of Illinois at Chicago | 2 |
| University of Oregon | 2 |
| University of Pennsylvania | 2 |
| University of Southern California | 2 |
| **University of Virginia** | 2 |
| Vanderbilt University | 2 |

Table 3: Universities with NSF AI Institute Awards. The left table is all of the universities that have been Primary (Lead) on an NSF AI Institute; the right table is all of the universities that have been subawardees on more than one NSF AI Institute. The 79 institutions with a single award as subawardee are not shown. The Virginia universities with one subaward are George Mason University are Norfolk State University. Other notable universities with a single subaward include Princeton University, Stanford University, and the University of North Carolina Chapel Hill.

Although ethics is often considered a branch of philosophy, here we take a broad view and include critical engagements with AI in other fields—from creative writing and history to education, psychology, and sociology—to the extent that these address the human impacts of AI from their respective disciplinary lenses and areas of concern. Regarding policy, we focus more narrowly on policies and practices operative within the university governing the practice of research, including those provided by the Office of the Provost.

#### 2.6.1 Survey Responses Addressing Ethics

The survey (Section 2.1) included two questions focused on ethics: (1) as a four-level response about the importance of ethical considerations to one's research, and (2) as a free text question asking for an elaboration on their response to the first.

In response the choice question, "How important are ethical considerations in your AI research?", the majority (58%) of respondents selected the top choice of *Important factor*, 13% indicated that they were *Starting to* *consider ethical issues*, 19% selected *No ethical considerations* (which may be depending on the nature of the research, or may indicate a lack of awareness of relevant ethical issues), and the remaining 10% selected *Unsure*. Although it is clear from these responses that most respondents are at least starting to consider ethical aspects of their work, the 29% of responses that indicated *No ethical considerations* or *Unsure* provide some justification for raising awareness about ethical concerns that apply to the preponderance of AI research.

In addition to a free text question asking to elaborate on that response, the survey asked respondents to describe any publications relating to ethics. Finally, the topics of ethics and policy in the context of AI research arose in other contexts, such as responses to resource needs. We summarize these responses in the aggregate, identifying broad areas of concern and themes, with the goal of describing the sample space of topics considered important by UVA researchers.

The following general areas of concern came across in these responses:

- *Bias and fairness*: Biases in data sets, algorithm designs, and development approaches may lead to unfair applications of AI technologies, especially when they are used in making critical decisions.
- *Privacy*: The rights of people to not have be surveilled without consent for the purposes of acquiring the data necessary to train models. There are legal and policy frameworks pertinent to collecting data on individuals, and developing rules and uncertain ethics about training models on different types of data, as well as uncertainty about what those models may reveal about their training data.
- *Repeatability*: The expectation that data acquisition and modeling processes be made available to the public, in the form of open source software and auditable activities.
- *Explainability and interpretability*: Addressing the problem of model opacity that plagues many machine learning methods, especially neural networks, where the logic behind a decision is not apparent from the trained model. When results are produced using an model that cannot be interpreted by its human trainers or users, there is a risk that model outputs are arbitrary or based on spurious factors.
- *Environmental impacts*: Training and executing large AI models can involve large amounts of energy as well as other resources.
- *Intellectual property and copyright*: The data used to train AI models, such as generative AI models of texts, images, and other media, often come from copyrighted sources and include content from individuals and organizations who did not consent to their use in model training. Whether such uses are covered by fair use is a disputed legal question, and even if such uses are determined to not violate copyright law that does not necessarily mean they are ethical.
- *Cybersecurity*: AI presents both new threats and opportunities for securing computing infrastructure, and AI models developed by researchers may be subject to unusual threats or may be usable for nefarious activities.

Survey respondents frequently mentioned “responsible AI” and “AI for the public good”, and expressed desire to do research that contributes to these goals.

In addition to the areas of concern, respondents also described three areas in which AI is having or is likely to have significant impacts which may be both positive and negative:

- *Economic*: AI will have impacts on business practices, labor markets, economic forecasting, and policy-making, and system-level phenomena. Examples where AI may have both positive and negative impacts include personalized pricing algorithms, the replacement and augmentation of work by AIs, and the automation of many activities previously requiring trained humans.

- *Social impacts*: Advances in AI and its deployment will have social effects beyond the narrowly economic, such as the application of AI to human decision-making in all social sectors, including government, law, medicine, and education.
- *Cultural impacts*: Related to social impacts are the impacts of AI on human language use, the production of art, and the pursuit of science.

In addition to identifying these areas, respondents provided several useful suggestions about how to address the need to develop and disseminate best practices. We synthesize these suggestions below:

- *Collaboration*: Instigate and support interdisciplinary research groups comprising core researchers (themselves from various disciplines), experts in ethics and policy, and, where applicable, members of the community.
- *Education*: Develop resources and venues to teach and inform researchers about the areas of concern described above. Vehicles described for this include workshops, courses, tutorials, and curated web resources.
- *Research*: Develop empirical research that examines AI from the design and development of technologies, to its use is educational and other settings, to its influence on high-stakes decision making. There was also support for research to investigate AI's impact on human behavior, relationships, and social structures.
- *Exploration*: Encourage artistic exploration in the form of installations and other works that explore and expose the effects of AI.

Finally, respondents urged a cautious and critical approach to support of AI research, and identifies needs for resources to become acquainted with best practices for the ethical use of AI in research, teaching, and service, as well as the need for clear guidance on government and university policies regulating research on AI.

#### 2.6.2 Interviews

Task force members conducted two group discussions on topics relating to ethics and policy: one with the Miller Center for Public Affairs at UVA, and one with a group of policy stewards at UVA.

**The Miller Center.** We interviewed the director (Marc Selverstone) and assistant director (Miles Efron) of the Miller Center of Public Affairs at UVA about their plans to apply AI to the corpus of presidential recordings curated by the Miller Center. The Miller Center's mission is to focus on presidential scholarship, public policy, and political history and to apply the lessons of history to the nation's most pressing contemporary governance challenges. The Miller Center has a 150-terabyte corpus of presidential recordings and discourse, covering presidents from Franklin D. Roosevelt to Ronald Reagan with a large collections of recordings from Richard Nixon and Lyndon B. Johnson.

Motivated by the need to transcribe, index, and explore major themes within these recordings, the Miller Center is planning a slate of high-impact AI projects including using large language models (LLMs), neural text embeddings, and multimodal AI to help researchers identify themes, tensions, and other patterns in its materials. We are excited by the prospect of applying AI to this work, both to advance the policy goals of the

Center, and to develop methods that that apply AI to humanistic research related to policy making.

**Policy Stewards at UVA.** To understand how UVA is considering the impact of AI on university policies, we conducted a group interview with Megan Lowe, Assistant Vice President and Chief of Staff; Brian Davis, Information Security Officer for Governance, Risk and Compliance; Kelly Hochstetler, Associate Vice President for Research Operations, Compliance & Policy; and David Hudson, Research Integrity Office. We were provided with a list of primary policies for review, with emphasis on RES-002, RES-009, and IRM-003, which cover the ownership, retention, and management of research records, data protection of university information, and the management of sponsored programs. Further discussion of these policies will help determine their relevance and application in the context of AI at UVA. Appendix E lists relevant documents and resources were identified by the group.

Currently, there are no specific policies addressing AI at UVA, but existing policies on data ownership, data use agreements, and general information security are being considered for adjustments to encompass AI applications. Ongoing discussions are considering establishing a governing body over AI to define constraints and guide policy development. Key areas of concern include copyright, domain-specific data regulation (including HIPAA which covers medical data and FERPA which covers educational data) with a focus on sensitive data storage, vendor requirements, privacy of input data, and licensing. Our conversations identified a need for transparency in AI usage and there are plans for task forces to extend policy into general AI and administrative areas.

Concerns have been raised regarding the potential constraints on research, the need for Institutional Review Board (IRB) considerations, training modules, data storage, sponsor agreements, and the transparency, trust, reproducibility, and explainability of AI models. Megan Lowe is working with the Provost's office to align policy with working groups, suggesting a coordinated approach to policy changes regarding AI. There are also ongoing discussions about incorporating AI into the definition of information technology and establishing guidelines for AI usage in non-research areas. There are also efforts underway to pilot a private ChatGPT instance, emphasizing data privacy by preventing data from being sent to externally-hosted LLMs.

#### 2.6.3 Analysis of Other Universities and Organizations

In addition to our local data gathering, we conducted a survey of how other universities are addressing the ethical and policy challenges of AI in relation to research. Appendix D provides information on organizations included in our survey. Although our survey was limited and cannot be claimed to fully represent all activities in this area, we can make some observations.

**Design vs. Regulation.** With regard to approaches to the ethics of AI, both in the context of research and more broadly, we find two kinds of engagements in the university setting: (1) efforts by AI users and developers to develop ethical AI models and methods, and (2) efforts by AI critics from a range of backgrounds to frame AI within larger social contexts. In seeking to integrate these perspectives, it is important to note that the different approaches represent potentially conflicting orientations and interests towards AI. They identify different areas of concern, assign risk differently, and propose different kinds of problem-solving mechanisms.

The first group, often under the name of *human-centered AI* (HCAI), originates within the core AI research community and tends to frame ethical problems in terms of safety, trust, and risk—terms that characterize relations between individuals and machines. This group focuses on product designs that integrate psychological, social and computational insights. In essence, it tends to define ethical problems in operational terms and seeks to avoids top-down legal and policy solutions by adopting technical solutions to mitigate ethical concerns. Concepts developed by the HCAI community include *alignment*, which considers how to ensure AI systems align with human values as intended; *augmentation*, with advocates for AI systems that enhance rather than replace human labor; *responsible AI*, which is used as a broad term encompassing ethics of data collection and use, and considering the impact and ethics of any use of AI; and *human-centeredness*, which refers to the general design perspective inherited from the established field of human-computer interaction (HCI) and user-centered design principles.

The second group, who typically embrace the term *AI ethics*, contains members who tend to come from the humanities and social sciences and focus on the systemic harms of AI on society as a whole. This group is more tightly aligned with philosophy and largely models itself after bioethics, a field that developed forty years ago in response to rapid technology changes in the medical space. In contrast to the first group, this group seeks to develop shared conceptual frameworks for reasoning about AI that will influence the development of laws and policies to regulate AI research and use. AI ethics inherits the vocabulary and conceptual framework of the academic field of ethics and established policy fields, such as law. For example, an AI ethicist may entertain consequentialist views of AI by adapting that body of knowledge to the new domain. This group has concentrated on defining the key areas of concern that characterize the domain of AI from an ethical perspective. These have been articulated in the *Rome Call for AI Ethics* (https://www.romecall.org/), produced by the RenAIssance Foundation, a non-profit organization established in 2021 by Pope Francis to promote anthropological and ethical reflection about the effects of new technologies on human life. This call, which has been signed on various AI ethics centers as well as Microsoft and IBM, identifies three impact areas—ethics, education, and rights—and six principles:

1. *Transparency*: AI systems must be understandable to all.
2. *Inclusion*: AI systems must not discriminate against anyone because every human being has equal dignity.
3. *Accountability*: There must always be someone who takes responsibility for what a machine does.
4. *Impartiality*: AI systems must not follow or create biases.
5. *Reliability*: AI must be reliable.
6. *Security and Privacy*: AI systems must be secure and respect the privacy of users.

These principles roughly correspond to the areas of concern identified by our respondents in Section 2.6.1.

## 2.7 Research Enterprise

Generative AI offers many potential capabilities that could be used to enhance the research enterprise and provide assistance to the activities of the VPR office. Currently, available generative AI tools provide a natural user interface through a question-and-answer format that accommodates multimodal data. With well-crafted prompts, generative AI models can produce text, graphs, tables, and executable programs, delivering answers quickly and with little effort. This would be a significant improvement over traditional analytics interfaces, which often require users to navigate through numerous steps and pages to generate plots and dashboards. The natural interface provided by generative AI enables users with varying levels of analytical proficiency to gain insights and produce analyses in flexible formats. Developing a UVA-specific model rather than depending on generic, publicly available models would position us to better understand and evaluate our research activities. This could result in a competitive advantage that allows our leadership to formulate precise, data-driven strategies for research support, keeping us ahead of our competitors.

The successful deployment of AI systems depends on three essential components: *well-defined user requirements*, *high-quality data*, and *appropriate AI models*. When it comes to AI models, a variety of techniques, including tuning large language models and Retrieval-Augmented Generation (RAG), have been demonstrated to be effective in data analytics applications. Many of these methods are accessible in open-source formats, and tuning, adapting, and executing such models is well within the computational resources and available expertise at UVA. Therefore, modeling does not constitute a hurdle to successful deployment, and we expect the state-of-the-art and what is available openly in these areas to continue to improve. However, the success or failure of an AI deployment project is also dependent on how well-understood and suitable user needs are for an AI solution and the quality and availability of the data needed to train and validate such a system. These two components are particularly context-specific, and our approach to evaluating the prospects for employing AI tools to assist the research enterprise at UVA focused on these aspects: assessing the current status of UVA's data infrastructure (Section 2.7.1) and identifying user requirements and application scenarios (Section 2.7.2). In addition, we provide a discussion of best practices at other institutions (Section 2.7.3).

### 2.7.1 Data Infrastructure

Generative AI and AI research enterprise tools fundamentally require integrated data to function effectively. These AI systems require access to diverse, yet interconnected, data sources to analyze patterns, predict outcomes, and generate insights that can significantly streamline research processes. The effectiveness of these tools is critically dependent on the quality and comprehensiveness of the underlying data.

Hence, to respond to the charge question about possibilities for employing AI tools in UVA's research enterprise, we conducted a thorough examination of the current maturity level of data infrastructure, the effectiveness of data governance practices, and the integration of legacy systems. These are the critical foundations for successful analytics and AI projects, so assessing this maturity level is crucial for identifying the extent to which the institution is equipped to support AI-driven analytics.

We gathered insights from the Office of the Vice President for Research (VPR) and the library department to gain a deeper understanding of these elements, collecting a detailed inventory of both internal and external data sources and analytics systems. This inventory includes evaluations of their functionalities, associated costs, and a quality assessment. We developed a table (available as the linked sheet: *Master Table for UVA Analytics*) that can serve as a blueprint for targeted investments in AI technology and infrastructure improvements. Our main observations from this analysis follow.

**Fragmented and incomplete data sources.** Because of the haphazard nature of data sources, simple queries often necessitate searches across multiple systems. The complexity of research activities at UVA, encompassing ongoing projects, proposals, Institutional Review Board (IRB) approvals, awards, sponsorships, and publications, cannot be fully captured by any single data source or system. For instance, Academic Analytics, while being one of the most extensive systems available, only includes information pertaining to UVA faculty, excluding research staff, postdoctoral fellow, and students. A preliminary case study focusing on a small cohort of faculty members exposed this system's lapses, notably the sporadic exclusion of certain publications from specific years. In contrast, the ResearchUVA (powered by Huron) system provides high quality data but suffers from functional limitations, such as the lack of collaborator network display and

ORCID integration. Consequently, a thorough search for a single researcher's activities often requires querying multiple systems such as searching for publications on Google Scholar, understanding collaboration networks on Academic Analytics, and looking up proposals on ResearchUVA.

**Inconsistent data.** Variation across data sources and systems prevents the discovery of a consistent information about UVA researchers. Attempts to cross-reference similar research activities across different platforms frequently results in conflicting outcomes. For example, a single researcher might have a drastically different number of publications and an H-index ranging from 13 to 57 across Google Scholar, Academic Analytics, and SCOPUS, thereby compounding the challenge of identifying a singular, reliable truth. If the outputs of myriad platforms are then combined and used to identify researchers, the results can be misleading when information from different sources is used in for different individuals and in different contexts.

**The absence of centralized data storage and data governance renders rapid AI development impractical.** AI development necessitates a centralized data infrastructure that regularly integrates data across internal and external systems. This infrastructure should be distinct from operational data storage, like Workday, which is designed to facilitate daily activities. From a technical perspective, a mature analytical system should incorporate both data warehouses and data lakes; however, UVA currently lacks deployment of these technologies. Another critical aspect is data governance, which includes determining data access controls, distinguishing between private and public data, and adhering to policy regulations. Presently, there are no codified policies or documentation concerning these governance issues. Collectively, these factors indicate that UVA's research enterprise is not ready to make a significant investment in AI development.

**Limited awareness and training on utilizing multiple systems can be potentially mitigated by natural language interfaces powered by AI models.** The need to consult multiple systems for answering even straightforward queries underscores the importance of users, particularly those involved in targeted analytics, becoming proficient with available tools and data sources. Feedback from discussions with a select group of users indicates a general lack of awareness about the breadth of available systems. Additionally, the varied interfaces across these systems pose a learning curve, necessitating training for users to perform tasks efficiently. Implementing a natural language-based interface like the ones provided by AI chatbots could potentially ease this challenge, but only with sufficient mechanisms to ensure results are reliable and users understand how to determine when they can be trusted. As a limited demonstration, Appendix F shows an example interaction with a prototype IRB chatbot that was created (in about 30 minutes of effort) by a task force member.

Based on these findings, Recommendation 16 emphasizes the need to improve underlying data availability and integration before any large-scale investments in AI implementation.

### 2.7.2 User Needs

Any successful technology deployment starts with understanding user needs, and AI tools are no different. Our goal in this analysis was to gain a better understanding of the needs of the research enterprise to understand the specific needs and challenges of stakeholders in the UVA offices, and to identify contexts in which AI could be applied beneficially. By aligning AI initiatives with the articulated needs of these stakeholders, the institution can ensure that the deployment of AI models yields significant, tangible benefits.

Considering the limited time available for this analysis, our process entailed engaging in dialogues with a wide array of stakeholders, encompassing leadership, faculty, and administrative staff. We followed the Gartner Analytics Enterprise Framework, which categorizes analytics stakeholders into three main types: *analytics consumers*, who utilize insights and reports to inform decision-making; *analytics producers*, who create data models, reports, and insights; and *analytics enablers*, who provide the necessary tools, technologies, data, and support data infrastructure to facilitate the analytics process. Accordingly, we organized our discussions and analysis of UVA's research enterprise around these three kinds of stakeholders, recognizing that many individuals act in all three roles.

**Lack of clear distinction between analytics producers and enablers leads to overload.** Creating data infrastructure and governance policies is a tremendous undertaking. The absence of a well-defined demarcation between those who produce analytics (the creators of data reports and insights) and those who enable analytics (the providers of data, tools, and infrastructure) often results in an undue burden on analytics producers. This overlap of responsibilities can lead to inefficiencies, bottlenecks in the production of analytics, and a strain on resources dedicated to generating actionable insights.

**Clear definitions of analytics tasks are needed.** Analytics tasks encompass a variety of complexity levels and time requirements, from real-time self-service analytics to in-depth analytical reporting. A clear design of the workflows and designated personnel will help analytics consumers understand when and where to obtain the needed resources. There is a pressing requirement for a structured framework that differentiates between real-time, self-service analytics and comprehensive, in-depth analytical reporting. Real-time analytics should empower users to access and interpret data on-demand without specialized skills, facilitating immediate decision-making. Conversely, in-depth analytical reporting should focus on providing detailed insights and strategic intelligence, requiring more time and expertise to produce.

**Unclear use cases.** Although there is a clear desire from analytics consumers to use AI to streamline and enhance decision-making, specific use cases where AI would be useful have not year been well identified and it is difficult to quantify the potential return-on-investment for any effort to employ AI tools in UVA's research enterprise. Based on our conversations with people in the VPR office, the most concrete scenarios identified were:

- *Trans-disciplinary research team assembly*: Facilitating the formation of competitive, cross-disciplinary research teams in response to specific funding opportunities, such as NSF solicitations, by identifying UVA researchers with the requisite expertise and interests and encouraging them to collaborate in pursuit of a particular opportunity. Any efforts in this direction need to be motivated by belief that faculty would be receptive to efforts to form teams this way and evidence that AI tools can usefully enable them better than simpler approaches.
- *Rapid research summaries for strategic conversations*: Providing instant, accurate information about specific research areas during critical meetings or conversations, such as enabling President Ryan to discuss research in an area of interest with prospective donors or stakeholders with informed confidence.
- *Expert identification for media engagements*: Quickly identifying and summarizing the qualifications of faculty members with expertise in niche areas for media appearances or interviews, thereby elevating UVA's presence in national discourse.
- *Strategic planning for future funding opportunities*: Analyzing current faculty strengths and identifying gaps in areas of expected growth in funding to inform strategic hiring and development plans that align with future funding trends.

Any investment in AI tools needs to be motivated by clear use cases and evidence that currently available tools such as public web searches, ResearchUVA (Huron), and reaching out to people directly would not provide better solutions than what is likely to result from a large investment in new tools.

#### 2.7.3 Case Study

The VPR brought to our attention an ongoing effort to use generative AI tools to support the research enterprise at the University of Pittsburgh. Here we report on what we have learned about that effort, and how it might inform UVA's decisions going forward. Our analysis is primarily based on notes from a conversation between a representative of the UVA VPR office and the Senior Vice Chancellor for Research at the University of Pittsburgh.[22] The University of Pittsburgh's Senior Vice Chancellor has a strong relationship with Microsoft and is leveraging this relationship in developing an AI-based tool. Although Microsoft does not plan to publicly release the product developed in collaboration with the University of Pittsburgh, if the pilot effort proves successful it could result in tools they are willing to develop with other universities. It is important to note that Microsoft's role and any product that results from this effort primarily fulfill the role of analytics producer. The value of any such product would depend on investments in internal data infrastructure necessary to facilitate the effective deployment of AI models.

We share similar desires for analytics solutions and issues in data infrastructure as those at the University of Pittsburgh. Like UVA, they face significant challenges primarily related to data access, privacy, and permissions as they work on integrating data from various sources, including both commercial platforms with proprietary frameworks and homegrown systems with differing levels of structure, into an AI platform. Issues include obtaining permission to use proprietary data, the complex process of integrating differently structured data, and ensuring appropriate access controls. The Vice Chancellor for Research, as the Chief Research Officer, can access all data but must navigate the complexities of data privacy (e.g., corporate sponsorships, intellectual property restrictions, national security restrictions, and HIPAA requirements for medical data). Further challenges arise in deciding how to share data with deans, department chairs, and faculty, which might involve creating access-limited versions of data or generating tailored reports, each requiring distinct tagging or framework setups. Moreover, incorporating reliable bibliometric data is problematic due to inconsistencies across sources and the high costs of potential solutions like those offered by Elsevier, compounded by the lack of standardized data collection by journals, such as the designation of contributors as faculty or students.

Despite similar data governance issues and a similar level of clarity of user requirements, the Senior Vice Chancellor at the University of Pittsburgh has a strong belief in the value of AI models despite uncertainty about their actual value. As successful effort to develop and employ AI tools would require cooperation and engagement from many individuals across the institution. Although there is instructive value in ambitious exploratory projects like the one currently underway at the University of Pittsburgh, the Task Force was not convinced that there is a clear use case or expected return on investment for such efforts based on our understanding of the needs of the research enterprise at UVA.

[22] Members of this Task Force were not invited to join the meeting, but were involved in discussions before it about what we wanted to learn about the effort at the University of Pittsburgh, and were provided with detailed notes from the conversation, which are the primary source for this analysis.

## 3 Advancing AI-Based Research at UVA

This section provides recommendations that respond to questions #1 and #5 from the Task Force charge:

- #1: *Hardware, software, data, human and organizational resources – Is the University providing adequate resources for UVA leadership/competitiveness in AI research across grounds?*
- #5: *What does UVA need to do to be considered competitive for programs like NSF AI Research Institutes or the USAISI?*

### 3.1 Computing Resources

> **Recommendation 1.** Provide predictable and sufficient long-term funding for Research Computing to develop and maintain long-term plans to support research activities.

The Research Computing (RC) organization has received occasional bursts of funding, but has suffered from the lack of certainty in its long-term sustainable funding. This uncertainty hampers effective budgeting and planning for upgrades, and means committed resources are not used as effectively as they could be.

The task force is encouraged by what we have heard about substantial and sustained funding commitments to support research computing at UVA, but it is unclear to us if what has been announced publicly so far as part of the Strategic Research Infrastructure Initiative[23] is sufficient, either in the certitude of long-term funding or the amount committed. We encourage the Provost to develop a funding model for Research Computing where a baseline level of funding is assured and can be used to support long-term positions and resource developments, and the level of funding is structured in a way that it will scale appropriately with increases in the overall research activity of the University.

> **Recommendation 2.** Develop a coherent model for supporting a mix of centralized and fully-shared computing resources along with distributed but centrally-managed and partially-shared computing resources.

Currently, investments in computing resources are done through an ad hoc mix of central, academic unit, and individual PI purchases. This has advantages in allowing autonomy for academic units and PIs, but results in substantial waste through duplication and underutilized resources. Individual academic units and PIs may have specialized needs and opportunities, and should be able to make local decisions about what equipment to purchase, but there should be a structure and incentives for equipment to be discoverable and partially shared. One incentive is the benefit of connecting the equipment to a larger infrastructure and benefiting from centralized technical support. In return for these services, the individual or unit funding the equipment would agree to a sharing model. The details of this will need to be worked out, but for computing resources that can be easily proportionally shared, we would suggest a model where the unit funding the resource would have priority for 75% of its cycles, and the remaining 25% as well as any cycles unused by the funding unity would

[23] https://research.virginia.edu/strategic-research/uva-research-update-october-19-2023

be available to the community. The Provost should encourage units that put a large amount of resources into their own computing resources to avoid unnecessary duplication and management overhead and transition to a centralized model, except in cases where local resources can be strongly justified.

**Recommendation 3.** Investigate opportunities for University-level commercial cloud contracts to improve individual researchers ability to use commercial resources.

Currently, many individual PIs end up needing to pay industrial cloud providers such as Amazon Web Services and Oracle Cloud for external compute resources as well as access to service APIs like those provided by OpenAI and Anthropic. When purchased by individual PIs, these resources can be very expensive and may be a limiting factor in resesarch. The University should track overall grant spending on industrial cloud computing, and consider opportunities for making agreements with providers to enable researchers to use these resources more efficiently.

### 3.2 Developing Other Resources

In addition to direct compute resources, we have several recommendations for other resources that should be developed for the community to support enabling AI-based research across the University.

**Recommendation 4.** Allocate research computing resources to provide beginners and non-experts with access to pre-installed, open-source generative AI models that encourage exploration without the need for a defined or funded project.

This recommendation addresses the gap between advanced users, who require specialized computing resources for complex projects, and beginners, who seek an accessible entry point for understanding the capabilities of AI tools and potential for incorporating them in their research. To bridge this divide, we propose creating an *AI Sandbox*, designed specifically for non-experts and those at the initial stages of AI learning. This would be a dedicated digital environment or service within UVA's research computing facilities. It would be equipped with pre-installed, open-source generative AI models, providing a setting where beginners can explore, experiment, and tinker with these models. This could start with a simple web interface, but also provide a path through computational notebooks and pre-built packages to more control and eventually lead to more advanced uses of AI tools.

To achieve this, we propose that Research Computing collaborates with current advanced AI users to select appropriate generative AI models and create accessible instructional materials. These materials should guide beginners through running basic models, interpreting results, and undertaking preliminary tuning exercises. Additionally, setting up ready-to-use, containerized environments can simplify the initial setup, allowing beginners to dive straight into AI practice without technical obstacles. This strategic approach could transform the *AI Sandbox* into a practical, hands-on learning tool within UVA's research computing framework, potentially serving as a foundational infrastructure for broader AI education initiatives.

### 3.3 Strategy

Before getting to our main recommendations about initiating an AI Reserach Center, we have one specific recommendation on financial policies, and two high-level strategic recommendations. The more concrete recommendations follow in the next section.

> **Recommendation 5.** If UVA wants to lead large-scale efforts, financial policies that discourage UVA from being the lead institution in a multi-university effort need to be avoided. The recently adopted research expenditures tax is particularly problematic and should be revised.

UVA has financial policies that work against our leading large-scale funding efforts. When teams of faculty from multiple institutions form plans for developing proposals and deciding on the lead institution, these costs are often a significant factor in deciding a different institution should lead the proposal. UVA should not have policies in place that put us at a disadvantage to peer universities, or expect PIs leading large projects to have the burden of negotiating special exemptions before being able to lead a competitive proposal.

In particular, the costs of managing a subcontract and the recent policy to tax research expenditures including those done through subcontracts make it undesirable for UVA to lead efforts such as AI Institutes. The expenditure tax on equipment purchases prevents UVA researchers from competing for large-scale infrastructure grants which would have negative financial impacts on their unit.

> **Recommendation 6.** Be *contrarian* and *creative*, or be *very rich*.

AI is an extremely expensive and competitive area, and most of our aspiring peer institutions have unsurprisingly also noticed its importance. To succeed, we need to either (1) make substantially larger investments than our competitors are making, or (2) do things that are different from what others are doing. If we can't outspend our competitors, UVA should select our investments strategically and focus on opportunities that are being underinvested in elsewhere and that take advantage of UVA's unique strengths. We will not be good or great by trying to do the same things everyone else is doing.

> **Recommendation 7.** Any redirection of centralized resources to support particular research areas, selected projects, or individual researchers needs to be strongly and carefully justified.

Although many of our recommendations encourage centralized investment in targeted activities, we think it is important to balance any such investments with the realization that overall resources are finite and any directed investment in one area takes away resources from other worthy areas. Any directed use of central funds should be supported by a justification that the investments are strategic and will have a better impact on advancing the University's mission than just providing "tuition rebates" to our students or overhead reductions to our researchers.

## 3.4 AI Research Center

Our main recommendations advocate for initiating and growing a pan-university center focused on AI research. We provide some details and specific suggestions for activities of the center in the following recommendations.

> **Recommendation 8.** Establish an AI Research Center at UVA with a mission to (1) advance, coordinate, and expand AI research at UVA; (2) foster interdisciplinary research efforts that advance research leveraging AI approaches and capabilities broadly across the university; (3) increase the visibility and impact of AI-based research at UVA.

The AI Center should have a Director, charged with leading the AI Center to accomplish its mission. Finding the right Director will be critical to the success of the center, and we encourage UVA's leadership to search broadly and creatively to find a Director, and to offer sufficient resources and empowerment in the position to make it attractive to an ambitious and talented leader. The Director needs to have awareness and appreciation for academic research, but does not necessarily need to be someone who has spent most of their career in academia or as a traditional researcher. Some understanding of AI is also important, but it is not essential for the Director to be a leading AI researcher. The Director should have a large amount of autonomy to lead the Center, but should be expected to provide regular updates on the activities of the center to University leadership.

It is essential that the Center is a pan-University effort and that all Schools have a significant role in some Center activities, as well as a stake in the success of the Center. One mechanism to ensure this would be to set up an advisory Board with representatives of all Schools of the University as well as other relevant organizations including Research Computing and the Library. Another mechanism is discussed in Recommendation 10. There should also be an external advisory board with members from industry, government, other universities, and alumni, to provide a broader perspective on the Center's activities.

Faculty affiliated with the center should benefit through access to opportunities and visibility, but should also be expected to contribute to the center — it should not be just a symbolic affiliation. Expected contributions would include acting as a mentor to an AI Fellow (see Recommendation 9), leading and participating in events at the Center including workshops and outreach talks, and providing periodic updates on their research activities for distribution through the center website.

## 3.5 Activities of the Center

The following recommendations elucidate expected activities of the AI Research Center, describing particular Center activities the Task Force believes would be valuable and should be supported through the Center. Although we provide suggestions for some specific activities and mechanisms, we emphasize that the Center Director should have a large amount of autonomy in determining the set of activities to pursue and how to implement them to achieve the Center's goals, with regular feedback and guidance from University leadership and the Center's advisory boards.

**Recommendation 9.** Initiate and support and *AI Fellows Program* to provide opportunities for (1) *AI Explorer Fellows*: researchers with ideas that could benefit from AI but without AI expertise to have opportunities to develop skills and understanding for using AI effectively in their research; and (2) *AI Ambassador Fellows*: researchers with AI expertise interested in training researchers in other fields and contributing to interdisciplinary projects that could benefit from AI expertise.

AI Fellows would be selected and coordinated through the AI Center using an open and competitive process, and would have appointments ranging from one semester to up to two years. AI Fellows will include researchers at various levels, including undergraduate students, graduate students, post-doctoral researchers, and both pre-tenure and post-tenure faculty. We would also encourage devising mechanisms to support visiting fellows who would be affiliated with the AI Center and contribute to projects while in positions in industry, government, and at other universities.

Student/postdoc level fellows will be provided with funding support and access to resources, as well as prestige, by being involved in the program. Senior-level and faculty fellows would be provided with funding support and release from other responsibilities.

To build a community among the fellows, enhance collaborations, and accelerate learning, the AI Center would organize regular meetings and events for the AI Fellows. All of the types of AI Fellows would also be expected to engage in outreach activities, such as workshops and public lectures to disseminate their research findings and promote AI literacy and ethics within and beyond the university.

Although it will be beneficial to provide seed funding to AI Fellows to support new projects, motivated by our Recommendation 7, any program that involves providing seed grants or other funding to individuals or small groups must be justified by strong arguments that this use of funds is a better way to advance the University's mission than other potential uses of those funds. Since this can be perceived as transferring overhead collected from nationally-competitive grants earned through the dedicated efforts by junior faculty to better funded and more senior researchers, making such an argument requires getting over many difficult hurdles:

- The work being funded this way is likely to contribute to the University's mission more than using the funds to support shared services and resources would.
- The work being funded this way could not be funded through available government funding agencies and other external sources for some good reason. This is tricky, especially in areas that are well supported at the national level like AI, since it requires explaining why the work is promising enough to be worth funding using UVA's funds but not good enough to receive competitive external funding.
- The selection process has a good likelihood of being both fair and capable of selecting the right bets to make. When selecting is done without using external reviewing, this is especially difficult because the most qualified people to technically review a proposal are usually excluded because of conflict rules.

The total costs of the program should be considered in this evaluation, not just the direct funding to be transferred. The total costs include the effort spent preparing proposals by both the winners and losers, which could otherwise have been spent on other activities including developing external proposals. They also include the costs of reviewing proposals for the selection process and in managing the awards after they are granted. Although much of this effort is done by faculty without any direct compensation, it is all taking time away from other activities.

**Recommendation 10.** Develop several AI Research Hubs, each dedicated to advancing AI research in a focused domain.

The proposed research center should establish focused research hubs, each dedicated to a different AI domain. By concentrating on specific, impactful areas such as AI Safety, AI in Healthcare, Foundational Research in Artificial General Intelligence, AI in Policy, and AI in Social Sciences, UVA can carve out niches where it leads regionally and globally, and can enhance its reputation and influence. Note that the preceding list of areas is intended as an illustrative list not a recommendation from the task force. Identify the right set of initial hubs and defining their foci will be an important early task for the Center, and will require the right balance of bottom-up faculty led efforts with top-down leadership-driven direction and selection to identify areas where UVA is well positioned to lead and where there is a core group of faculty committed to building a successful research hub. The AI Center should be broad and inclusive, but the effectiveness of the research hubs will depend on them being sharply focused.

Importantly, to foster a truly interdisciplinary and university-wide initiative, leadership roles in the AI hubs should be spread across the Schools of the University, ensuring a comprehensive, collaborative, and diversified approach to AI research and application. The AI Fellows would be involved in relevant AI Research hubs, but the hubs would also provide a mechanism for coordinating research across the University and for involving other faculty and students in center activities.

One activity of the research hubs could be to pose specific challenges to focus multi-disciplinary research. These challenges would target hard unsolved problems across various disciplines (e.g., autonomous systems, health, environment, democracy). This would not only position UVA as a center for cutting-edge AI research but also foster collaboration with top institutions worldwide. UVA has already laid the groundwork in several of these domains through initiatives like precision health, neuroscience, and democracy under its Grand Challenges program. A seamless and strategic approach would integrate existing initiatives under the umbrella of AI Research Hubs, while continuing to nurture other promising research areas.

**Recommendation 11.** Provide centralized support for initiating and maintaining data consortia.

AI programs are as good as the data that supports them. In the cases where UVA is a national leader for a specific type of data we will create consortia for use of that data. For example, the School of Medicine and Medical Center is a national leadership role in the use of Physiological Monitoring data. This involves both ground up support, i.e. reserved time to participate in the consortium, and top down support, i.e. focused leadership to sustain the work and lead the pursuit of opportunities. Understanding and access to data is always a underlying requirement for advancing in AI. The Center for Advanced Medical Analytics (CAMA) and the DOM Clinical Analytics Lab (CAL) are two existing examples of data domain groups. Although much of the effort required to build a successful data consortium depends on individual PIs with data and connections, some of the effort to start such a project is duplicated across every data consortium, and the long-term success of a data consortium depends on institutional support and commitment.

> **Recommendation 12.** Develop innovative programs to incentivize and encourage formation of effective teams for interdisciplinary research involving AI.

Designing such programs is challenging and would be primarily up to the Center Director and their team. As a starting point for discussion, motivated by both the appeal and the failure to meet its goals of the 3CAVS program, we recommend developing a program with similar aims but a flipped mechanism:

**SVAC$^3$ (Spurring Virginia's Amazing Collaborators to Create Chaotically) Program**

1. Researchers interested submit interest form, indicating their own areas of expertise, some types of people they would be interested in working with, and schedule availability, and a list of people they already collaborate with.
2. Organizers use an algorithm to form 3-person teams with best-effort matching.
3. Emails are sent to the team to initiate a group.
4. Teams formed in this way have three weeks to come up with an idea and submit a short proposal for what they would work on together.
5. From the submissions, select as many teams as possible for funding at a level that is enough to at least fund a full student for 2 years, with options to extend and expand if initial results are promising. The funding should have as few restrictions as possible on both how it can be spent and when it expires.

The success rate of such a program should be low—most randomly arranged teams will not yield a successful collaboration. But, when formed teams do find worthwhile opportunities these are ones unlikely to be found otherwise, and the initial effort is low and the ancillary benefits are high. Getting the details of how to do the matching and how to structure the program right is difficult, and its important the leaders initiating such a program realize that it may take a few attempts over multiple years to get things working as intended, but once a successful model is developed it will be a sustained pipeline for building interdisciplinary collaborations and enhancing the research environment at UVA.

# 4 Ethics and Policies

This section addresses charges 2 and 3 from the committee charge:

- #2: *Should the University be providing training for all UVA researchers in AI ethics, fairness, bias and related areas? Make recommendations about the type and modes of training.*
- #3: *Do UVA policies on research and data need to be updated due to AI?*

In combining these, we address the topics of ethics and policy in relation to AI research, broadly conceived. Although these comprise distinct areas of expertise, we combine them as both are related to the general concept of value and the evaluation of the positive and negative impacts of technology on human well-being. Whereas ethics focuses on framing and understanding the risks and harms associated with technologies, policy concerns shared and codified rules, both internal to the university and legal, created to govern the use of AI consistent with human value.

**Recommendation 13.** It is not necessary or appropriate at this time to develop any UVA-specific training materials or to require any specific AI-related ethics training. Instead, UVA should endeavor to raise awareness of ethical issues connected with AI-based research and provide well curated resources to help researchers.

To enhance AI ethics practices within the research community, we recommend adopting a holistic and process-oriented approach that emphasizes continuous education, awareness, and support to foster an environment of ethical vigilance and responsible research conduct.

A small team should be formed that is responsible for curating and maintaining an accessible repository of resources pertinent to AI research including relevant laws, policies, regulations, and guidelines. Resources should be tailored to accommodate varying levels of engagement, from simple links for quick reference to comprehensive handbooks for in-depth understanding. The curation process should be inclusive, non-judgmental, and designed to provide a broad perspective, facilitating informed decision-making rather than prescribing practices (other than where established laws apply). The resources should be periodically updated, ensuring that UVA's resources remain abreast of the latest legal and ethical developments in AI. In addition, this team should also be knowledgeable about the potential risks associated with specific research activities and able to offer informal guidance to UVA reseachers.

**Recommendation 14.** Relevant legal regulations and policies of granting agencies and publisher should be summarized and telegraphed to researchers to ensure compliance.

Many federal granting agencies and academic societies and publishers have or are developing their own guidelines for generative AI use, and it is important the UVA researchers are aware of these policies and understand how to comply with them. We encourage the VPR office to provide resources and communicate with researchers in a manner that makes them aware of exigent requirements in the conduct of research. In most cases, the closer the communications are to where the researcher needs to be aware of issues the better. Forms of delivery may include a website with (appropriately qualified) executive summarizes of important laws and policies, as well as checklists provided for specific use cases, such as the training of language models in the context of human subjects research.

**Recommendation 15.** Policy stewards should consider the implications of AI when developing and revising UVA policies.

New technologies are rarely a reason for a fundamental policy change, but technological change may require policies to be stated with clearer and more explicit language. There may also be unique situations arising from the use of AI that may require new language and, possibly, new policies. Emerging situations include those arising from the relationship between models and data introduced by training process where the amount of information a model reveals about its training data, and the implications of distributing the model in different ways, are not yet clearly understood. In these cases, the ethical and legal relationship between data producers and data users in the context of research may need to be clearly addressed.

In any place where AI models are considered for decision-making processes, it is important that potential biases in the models are evaluated and mitigated. In addition, if outputs of AI models have an impact on critical decisions, it is important to document how those models were developed and tested, and provide explanations for their predictions.

# 5 AI for the Research Enterprise

Charge question #4 concerns the evaluation of the *feasibility* and *inherent value* of employing AI tools to support various aspects of the research enterprise under the purview of the VPR office at UVA:

> #4: *Use of AI to facilitate the UVA research enterprise – should VPR research development start to develop AI models to be able to mine our own data (papers, submitted grants, IRB protocols, animal protocols, IP, etc) in order to optimize research development at UVA? Others are already doing this with impressive results, and we may fall behind without such an effort.*

We have two main recommendations responding to this charge question.

> **Recommendation 16.** Before investing in a large-scale AI enterprise implementation, UVA should focus on improving research data integration. In addition, before any major investments are made in AI for the research enterprise, the VPR office needs to identify use cases where AI may be of value.

The success of an AI application fundamentally relies on three key elements: high-quality data, well-designed user experiences, and accurate models. Our analysis in Section 2.7.1 and Section 2.7.2 finds that we currently fall short in the first two areas. Bridging these gaps requires significant effort to integrate various data sources and improve the quality of data available. Doing this is both a precursor to any major investment in AI tools and an effort that would have other significant benefits to the operations of the VPR office.

Before any major investment is made in AI tools for the research enterprise, we advocate for a limited pilot study to clarify user requirements and demonstrate value in an area with relatively high-quality data. After completing the pilot study, we should reassess the feasibility and return on investment of developing AI-based tools for the research enterprise before making significant investments. It is important to understand how the outputs from any such tools would be used, and what impact their use would have on UVA's operations.

To begin this process, the VPR office should identify and document a series of clear, compelling use cases where AI tools could significantly improve the efficiency and effectiveness of accomplishing their mission and advancing research at UVA. If any promising use cases are identified, the next step would be for the VPR office to initiate a focused, small-scale AI research enterprise pilot, centered on a selected use case. This pilot may be conducted within a single school that has a cohesive data structure or limited to a segment of readily accessible high-quality data. Its goal is to gain a deep understanding of user requirements, develop a robust technical framework, and showcase the potential benefits of AI tools in research and administrative efficiency. This serves as a crucial experiment to evaluate the practical benefits of AI tools and identify any potential issues in a controlled setting before making the substantial investment required for a broader implementation.

High-quality, well-integrated data is crucial for the success of future AI initiatives. Hence, our second recommendation in this section emphasizes the importance of UVA's continued investment in data integration efforts and the need for focused leadership of these efforts:

**Recommendation 17.** Establish a dedicated University-level Data Steward role.

As highlighted in Section 2.7.1, UVA's research infrastructure currently faces significant challenges in data integration, which is essential for the successful deployment of AI applications. Essential data elements, including research proposals, grants, expenditures, Institutional Review Board (IRB) protocols, patents, and publications, are dispersed across various systems with limited interconnectivity. This fragmentation hampers not only the potential for advanced AI-driven analyses but also the basic functionalities of descriptive analytics and interrogation.

As discussed in Section 2.7.2, the current team structure lacks a distinct separation between analytics enablers, responsible for data management, and analytics producers, such as analysts and data scientists. Frequently, analytics producers find themselves juggling both roles, which results in inefficiencies within the system. This challenge is not unique to UVA (Section 2.7.3), but is prevalent at most organizations, and addressing this well may provide advantages to UVA.

We recommend establishing a dedicated **Data Steward for Data Integration and Analytics** role to manage this effort. Specifically, this role is dedicated to overseeing the organization's data integration efforts, developing and maintaining analytical data repositories, and ensuring data quality and accessibility to empower analytics and AI initiatives. The ideal candidate will possess a strong foundation in data management practices, technical proficiency, and an understanding of data governance principles to facilitate data-driven decision-making processes across the organization.

Key responsibilities for the Data Steward role include:

- *Data integration and management*: Develop strategies for integrating data from diverse sources to create a unified data environment, designing and maintaining data pipelines, and ensuring adherence to governance standards.
- *Analytical data storage development*: Design and manage analytical data stores (data warehouses or data lakes), oversee the data processing workflow, and implement data modeling to support analysis and modeling efforts.
- *Data quality and governance*: Establish data quality standards, maintain data dictionaries and metadata repositories, and collaborate on defining data policies and roles.
- *Stakeholder collaboration and support*: Engage with business analysts, data scientists, and other stakeholders to meet their data needs, offer expertise on data-related issues, and promote data literacy and governance.
- *Continual innovation*: Keep up-to-date with emerging data management, analytics, and AI technologies, and recommend tools and technologies to enhance the organization's data infrastructure.

# 6 Other Recommendations

Our final charge was,

> #6: *Anything else the committee brings to light.*

Following this license, we have two recommendations outside the specific scope of our charge, but which are based on issues that came up repeatedly in the committee's discussions.

> **Recommendation 18.** If we want to change the incentives for early career researchers, we need to reduce the emphasis on external letters in the Promotion & Tenure process.

During our committee discussions as well as our Forums, it came up several times that there was a need to change the incentives in some academic units if we want to encourage early-career faculty to tackle ambitious and risky projects, do interdisciplinary work, and invest time and effort in building and sharing research tools. Although the University can make pronouncements about valuing different types of work and even put encouraging statements into our Promotion & Tenure guidelines, any efforts from leadership here are futile as long as early career researchers have the expectation that external review letters will be the primary factor in determining their promotion. Candidates know the reality that people writing those external letters will be unlikely to read or be influenced by any unusual aspects of UVA's P&T guidelines.

The only way UVA's leadership can significantly change the incentives and pressures that influence the decisions of tenure-track faculty is if we reduce the reliance on external reviews and place more trust in our internal evaluations. This would be a bold and controversial action, and must be done in a way that does not adversely impact faculty at different stages in their current careers. Any increase in reliance on internal evaluations should also be accompanied with increased clarity and transparency about the criteria that will be used in evaluating candidates, as well as faculty involvement in establishing those criteria.

A lighter way to reduce the emphasis on external letters would be to simply decrease the number of external letters expected and liberalize the guidelines for how external reviewers are selected. The Provost's review could impose a maximum number of letters permitted, and hope this would propagate to influencing school policies. The current policy[24] seems to not specify number of external letters for normal process, but that for an expedited review, "Three outside, arms-length letters, are acceptable". We would urge a similar policy at all levels, and a stronger statement that would only allow cases to include a maximum of three external letters, and guidelines that encourage nominations where all external reviewers are selected from a list provided by the candidate.

> **Recommendation 19.** Be more transparent and communicate operations, actions, and plans more effectively with faculty and the broader community.

Administrative leaders should communicate more effectively with the faculty and broader community, and be

[24] https://uvapolicy.virginia.edu/policy/prov-017

more willing to share information about funding, organizational efforts, and plans. Three specific examples we encountered carrying out the charge for this task force are:

- *Transparency about research overhead distribution*: Nearly all of the research overhead at the University is generated by the efforts of faculty, and a substantial fraction of funds raised through these efforts is captured through overhead charges. Despite this, the typical faculty member's understanding of how research overhead works is limited to a cartoonish and probably grossly distorted view: overhead rates always go up, expected services always decline, no one has any idea where the money goes or what is done with it, and only very well-connected faculty can negotiate special deals to pay less. We are sure the reality is different, but until the administration shares information about how overhead is collected and used more openly most faculty will be left guessing and likely assuming the worst. Part of this transparency would also include openness about the budgets of administrative units, including the VPR office, and communicating this information clearly in easy to find places on appropriate websites.
- *Transparency about investments*: As members of this task force, we have been privy to some details on previous and planned investments in research by the University, including the actual size in dollars of certain investments that have not been shared publicly. Any investments the University leadership makes should be ones where both the nature and scale of those investments can be justified as contributing to our missing in a way that has a larger expected return, and making the UVA community and external world (especially those we hope will join us, or work with us) aware of these investments has tremendous benefits.
- *Clear communication about ad hoc committees and task forces*: During our forums and other discussions, we were often asked if there was a website or other document with more information about our task force and had to respond that no such information existed. Some of the task forces established by the Provost's Office are publicly announced, others (like ours) are not announced, and there is no public web pages that lists all of the efforts underway. Through our involvement in this task force, we have become aware of at least six different University-level groups charged with making recommendations on various aspects the impact of Artificial Intelligence on the University, but there is no public web page that describes all of the groups, their charges, and how community members should engage with them.

In general, we advocate for transparency and open communication about all Provost office activities except in cases where there are strong legal, personal, or political reasons to keep things confidential.

# 7 Summary of Recommendations

Here we recap all of our nineteen recommendations, but instead of ordering by the study charge questions, organized around the task force's view of what would be required to implement them.

*Recommendations that urge cautious steps, limited immediate investments, and exploratory steps:*

- Recommendation 3: Investigate opportunities for University-level commercial cloud contracts to improve individual researchers ability to use commercial resources.
- Recommendation 13: It is not necessary or appropriate at this time to develop any UVA-specific training materials or to require any specific AI-related ethics training. Instead, UVA should endeavor to raise awareness of ethical issues connected with AI-based research and provide well curated resources to help researchers.
- Recommendation 14: Relevant legal regulations and policies of granting agencies and publisher should be summarized and telegraphed to researchers to ensure compliance.
- Recommendation 15: Policy stewards should consider the implications of AI when developing and revising UVA policies.
- Recommendation 16: Before investing in a large-scale AI enterprise implementation, UVA should focus on improving research data integration. In addition, before any major investments are made in AI for the research enterprise, the VPR office needs to identify use cases where AI may be of value.

*Recommendations for immediate policy changes:*

- Recommendation 5: If UVA wants to lead large-scale efforts, financial policies that discourage UVA from being the lead institution in a multi-university effort need to be avoided. The recently adopted research expenditures tax is particularly problematic and should be revised.

*Recommendations requiring cultural changes and bold and effective leadership, but not significant resources:*

- Recommendation 2: Develop a coherent model for supporting a mix of centralized and fully-shared computing resources along with distributed but centrally-managed and partially-shared computing resources.
- Recommendation 6: Be *contrarian* and *creative*, or be *very rich*.
- Recommendation 7: Any redirection of centralized resources to support particular research areas, selected projects, or individual researchers needs to be strongly and carefully justified.
- Recommendation 18: If we want to change the incentives for early career researchers, we need to reduce the emphasis on external letters in the Promotion & Tenure process.
- Recommendation 19: Be more transparent and communicate operations, actions, and plans more effectively with faculty and the broader community.

*Recommendations requiring significant investment:*

- Recommendation 1: Provide predictable and sufficient long-term funding for Research Computing to develop and maintain long-term plans to support research activities.
- Recommendation 4: Allocate research computing resources to provide beginners and non-experts with

access to pre-installed, open-source generative AI models that encourage exploration without the need for a defined or funded project.

- Recommendation 11: Provide centralized support for initiating and maintaining data consortia.
- Recommendation 12: Develop innovative programs to incentivize and encourage formation of effective teams for interdisciplinary research involving AI.
- Recommendation 17: Establish a dedicated University-level Data Steward role.

*Recommendations requiring substantial investment and buy-in from leadership across the University:*

- Recommendation 8: Establish an AI Research Center at UVA with a mission to (1) advance, coordinate, and expand AI research at UVA; (2) foster interdisciplinary research efforts that advance research leveraging AI approaches and capabilities broadly across the university; (3) increase the visibility and impact of AI-based research at UVA.
- Recommendation 9: Initiate and support and *AI Fellows Program* to provide opportunities for (1) *AI Explorer Fellows*: researchers with ideas that could benefit from AI but without AI expertise to have opportunities to develop skills and understanding for using AI effectively in their research; and (2) *AI Ambassador Fellows*: researchers with AI expertise interested in training researchers in other fields and contributing to interdisciplinary projects that could benefit from AI expertise.
- Recommendation 10: Develop several AI Research Hubs, each dedicated to advancing AI research in a focused domain.

# A Data and Availability

This appendix provides information on sources of data for this report that are not described elsewhere.

The AI-generated image of the committee inside the cover was produced by StableDiffusionXL using the prompt:

> 8 people meeting to discuss recommendations for a report, University of Virginia, style of paintings of Salvador Dali, the picture should show 8 committee members and no other people. The room should be in the style of a conference room in a library. The committee is working very hard, and focused on the task.

Table 3 and the discussion in Section 2.5 is mostly based on data from https://aiinstitutes.org, which scraped and analyzed using custom code to produce Table 3. The code is available in https://github.com/AIResearchTaskForce/nsfaiinstitutes.

The code for performing the topic analysis on the OpenAlex data and the bibliometric analysis of the Scopus data is available in https://github.com/AIResearchTaskForce/bibliometrics.

# B Survey Details and Further Analysis

This appendix provides more details and analysis on the survey described in Section 2.1.

## B.1 Trends and Correlations in AI Use and Understanding

Table 4 shows the relative frequency of response levels for each question. For example, $41\%$ of respondents answered that they have *never* used AI in their research to the frequency question, while $46\%$ answered that they expect to do so in the expectation facet. Table 7 shows correlations among these aggregate response rates.

These data exhibit some general trends.

First, **centrality** and **frequency** are highly correlated ($r = .91$), and are both skewed right, with relatively low mean levels of $1.24$ and $1.31$ respectively (see Table 5). In each case, level $0$ accounts for the majority of responses, while the other levels are more evenly distributed. We may infer that in the sample population a significant number of people never actually use AI in their research ($frequency = 0$ ), and that a similar proportion consider it of minor significance to their work ($centrality = 0$). Whether or not these are the same people may be inferred from the correlation of non-aggregated facets in the survey data; these are presented in Table 6. In fact, the two facets are strongly correlated among users ($r = .85$). Thus, we may infer that frequency and centrality are mutually predictive. We may also surmise that there are two camps among those interested in AI—those who *use* it and those who *observe* and critique it. Taking the centrality measure as an index of this division, we suggest a $60/40$ split among users and critics in the sample group, where it is understood that the critics are more clearly defined in their response level.

Second, there is a very high aggregate correlation between **comfort** and **potential** ($r = .92$). For these facets, the responses are skewed more left, with a mean level of $2.13$ and $2.27$ respectively. Thus, a majority of respondents are comfortable with AI, and a similar majority see its potential for their research. This

Table 4: Response rates per facet answer.

| level<br>facet | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| centrality | **0.51** | 0.12 | 0.10 | 0.14 | 0.12 |
| comfort | 0.12 | 0.16 | **0.30** | **0.30** | 0.11 |
| expectation | **0.27** | 0.01 | 0.17 | **0.46** | 0.08 |
| familiarity | 0.06 | **0.34** | **0.29** | 0.22 | 0.08 |
| frequency | **0.41** | 0.16 | 0.20 | 0.16 | 0.07 |
| potential | 0.11 | 0.15 | **0.26** | **0.29** | 0.18 |

Table 5: Mean levels per facet.

| facet | mean level |
|---|---|
| potential | 2.27 |
| comfort | 2.13 |
| expectation | 2.05 |
| familiarity | 1.92 |
| frequency | 1.31 |
| centrality | 1.24 |

would suggest that those who are comfortable using AI are also those who see the most potential for it in their research. However, this is not confirmed by the moderate non-aggregate correlation of the two facets ($r = .55$). In other words, individuals with a level 2 or 3 comfort with AI are not necessarily those who see its potential for their research. This suggests that there are many who are not comfortable with AI but who do see its potential, as well as many who *are* comfortable with AI but *do not* see its potential. Thus, comfort and potential are not mutually predictive, even though they are both relatively high.

The third trend is the high number of respondents—$46\%$—who have relatively high **expectations** for increasing the use of AI in their research ($expectation = 3$), although only $8\%$ expect the main focus of their future projects to be AI ($expectation = 4$). Interestingly, the relationship between expectation and potential is opposite to that of comfort and potential. While the aggregate correlation between **expectation** and **potential** is moderate ($r = .50$) , the non-aggregate correlation is high ($r = .72$). *This suggests that those who expect to use AI in their research do so because they perceive its potential, not because they are comfortable with it.* Expectation also has an interesting relation to **centrality** and **frequency**. Whereas the latter suggest a high percentage of respondents with low levels (non-central and infrequent), the former has equally high numbers for Level 3 (although not for Level 4). This suggests a split between two groups in the sample—those who have no plans to use AI and those who do, and, within the latter, smaller group who plan to make it central to their work ($expectation = 4$). This corroborates the suggestion made above that the response group may be divided into critics and users.

Table 6: Facet correlations (top).

| facet 1 | facet 2 | corr ($r$) |
|---|---|---|
| centrality | frequency | 0.85 |
| expectation | potential | 0.72 |
| frequency | potential | 0.71 |
| centrality | potential | 0.69 |
| familiarity | frequency | 0.68 |
| centrality | familiarity | 0.68 |
| comfort | frequency | 0.63 |
| expectation | frequency | 0.63 |

Table 7: Aggregate facet correlations (all).

| facet 1 | facet 2 | corr ($r$) |
|---|---|---|
| comfort | potential | 0.92 |
| centrality | frequency | 0.91 |
| comfort | familiarity | 0.56 |
| comfort | expectation | 0.55 |
| expectation | potential | 0.50 |
| familiarity | potential | 0.42 |
| expectation | frequency | 0.31 |
| centrality | expectation | 0.29 |
| comfort | frequency | -0.16 |
| expectation | familiarity | -0.22 |
| familiarity | frequency | -0.31 |
| centrality | comfort | -0.44 |
| frequency | potential | -0.47 |
| centrality | potential | -0.63 |
| centrality | familiarity | -0.64 |

## B.2 Use of AI in Research (Facets)

The survey asked respondents to answer to a series of questions regarding facets of their relationship to AI in the context of their research. A subset of these questions were presented in the form of multiple choice answers that correspond to increasing levels of commitment. Although not explicitly presented as five-level Likert scales, here we interpret them as such, using a scale of 0 to 4, where 0 signifies the lowest level value and 4 the highest. These questions address the following facets:

- **Familiarity** level with AI in general.
- **Comfort** level in using AI tools.
- **Frequency** of AI use in research projects.
- **Centrality** of AI to projects that use it.
- **Expectation** for Artificial Intelligence Usage in Future Projects.
- **Potential** of Artificial Intelligence in Enhancing Future Research.

Descriptions of the five levels for each facet are presented in Table 8.

### B.2.1 Response Rates

Although all respondents answered at least one of the Likert-type questions, not all such questions were answered by all respondents. For all but one question—concerning the centrality of AI in their research—the response rate hovers around 90% (see Table 9). It is reasonable to assume that the non-responses to the centrality question (which has around a 40% smaller response rate) are effectively 0s, i.e. they indicate ”minor significance,” since respondents are likely to answer this question if, in fact, they find AI in any way significant to their research. Therefore, to compute statistics from the data, we impute a 0 for the missing centrality responses. However, we drop observations that have non-responses to any of the other questions. This yields a sample size 204, or 88% of the response group and 13% of the faculty population.

### B.2.2 Free Text Responses: Specific Topics

- AI for science
- AI in artistic production
- Preliminary unpublished data on proteins structure
- AI in information technology
- AI for quantum computing
- Explainable AI
- AI in creative fields
- AI hardware
- Large Language Models
- Signal processing
- AI in popular culture
- Knowledge representation, ontologies, and provenance awareness in AI
- Foundations of pre-model explainability

Table 8: Facet response types.

| | |
|---|---|
| **Familiarity with Artificial Intelligence (familiarity)** | |
| 0 | Not familiar at all |
| 1 | Somewhat familiar |
| 2 | Moderately familiar |
| 3 | Very familiar |
| 4 | Expert |
| **Comfort Level in Using Artificial Intelligence Tools (comfort)** | |
| 0 | Very uncomfortable |
| 1 | Uncomfortable |
| 2 | Neutral |
| 3 | Comfortable |
| 4 | Very comfortable |
| **Frequency of Artificial Intelligence Usage in Research Projects (frequency)** | |
| 0 | Never |
| 1 | Rarely |
| 2 | Sometimes |
| 3 | Often |
| 4 | Nearly Always |
| **For your projects that use Artificial Intelligence, how central is it to the project (centrality)** | |
| 0 | Minor significance |
| 1 | AI is helpful and improves efficiency, but not necessary |
| 2 | AI is an important tool, but if it were not available the project would continue |
| 3 | AI is essential to the project, which could not be done without it |
| 4 | The primary goal of the project is to advance AI |
| **Expectation for Artificial Intelligence Usage in Future Projects (expectation)** | |
| 0 | No future use of AI |
| 1 | Reduce from current usage* |
| 2 | Maintain current usage |
| 3 | Increase usage of AI |
| 4 | Expect main focus of future projects to be AI |
| **Perceived Potential of Artificial Intelligence in Enhancing Future Research (potential)** | |
| 0 | No potential |
| 1 | Low potential |
| 2 | Moderate potential |
| 3 | High potential |
| 4 | Very high potential |

* This answer does not clearly align with a Likert scale.

Table 9: Response rates per facet.

| facet | response rate |
|---|---|
| familiarity | 0.93 |
| comfort | 0.93 |
| frequency | 0.93 |
| centrality | **0.55** |
| expectation | 0.88 |
| potential | 0.88 |

- Smart manufacturing
- Semiconductor manufacturing

#### B.2.3 Selected Publication Summaries

*Briefly describe one published research paper in which you used Artificial Intelligence methods/techniques/models.*

1. **Core Machine Learning and Responsible AI:** Research focuses on core ML with subtopics like responsible AI, emphasizing privacy, fairness, and robustness, and the development of ML models to satisfy verifiable properties.
2. **Physics and Science:** Applications in physics are evident, such as using ML to study neutrino oscillation parameters and analyzing literature data on material growth, highlighting interdisciplinary collaboration.
3. **Economics and Market Analysis:** ML models like BART are applied to study market dynamics, for instance, analyzing prices and profits in the antiquities market.
4. **Legal and Social Science:** ML techniques help examine verbal expressions of confidence in legal contexts, adding unique value in understanding eyewitness identification accuracy.
5. **Healthcare and Biology:** AI is employed in various healthcare applications, from decoding brain activity patterns to image segmentation and classifiers for medical imaging, demonstrating AI's impact on diagnostics and treatment planning.
6. **Environmental and Earth Sciences:** AI techniques, including reinforcement learning, are used for designing water systems control policies and forecasting natural events, showcasing AI's role in environmental management and disaster preparedness.
7. **Engineering and Technology:** Research includes using AI for robustness in classifiers, artifact subtraction in imaging, and developing neural networks for specific engineering tasks, reflecting AI's integration into technical innovation.
8. **Social Media and Online Communities:** AI methods like NLP and topic modeling are utilized to understand communication patterns and online community dynamics, illustrating AI's applicability in digital social research.
9. **Interdisciplinary Applications:** The response encapsulates a diverse range of AI applications across fields, indicating a broad integration of AI methods in various research domains, from theoretical studies to practical implementations.

### B.3 Research Interests

Responds were asked to describe their research interests from three perspectives: (1) work in progress, (2) research goals, and (3) a possible project. The second two questions focus on aspirations, which provide a balance to the first question as well as the information acquired from the questions on publication.

#### B.3.1 Research in Progress

> Briefly describe one work in progress (paper/project) in which Artificial Intelligence methods/techniques/models is used (optional - please only share high level information about any unpublished work).

1. **Event Reconstruction:** Utilizing AI in physics for event reconstruction in experiments like Mu2e.
2. **Patient Comorbidities:** Applying large language models to predict patient comorbidities from clinical notes.
3. **Reinforcement Learning in Systems Control:** Exploring reinforcement learning for stormwater system control and hydropower system adaptation to climate change.
4. **AI in Decision-Making:** Investigating how AI can improve human decision-making in contexts like eyewitness identification accuracy.
5. **Molecular Dynamics:** Employing AI to study peptide assembly through molecular dynamics.
6. **AI in Education:** Developing AI models to classify instructional activities in classroom videos.
7. **Medical Imaging:** Using deep learning to correct imperfections in magnetic resonance images and for protein interaction prediction in ultrasound imaging.
8. **Ecology and Evolution:** Applying machine learning to predict factors associated with population extinction and colonization.
9. **Quantum Computing:** Integrating AI with quantum computing for physics simulations.
10. **Cyber Defense:** Developing reinforcement learning agents for cyber defense.
11. **AI in Health Ethics:** Calling for more AI expertise in health ethics and data governance.
12. **AI in Pricing Research:** Conducting research at the intersection of machine learning and pricing.
13. **Legal Text Analysis:** Using natural language processing to examine judicial judgments and their social impacts.
14. **Mechanical Property Estimation:** Utilizing imaging signatures to estimate tissue mechanical properties.
15. **Social Impact of AI Decisions:** Studying how AI-based decisions influence prosocial behavior.
16. **AI in Language Processing:** Developing better language understanding algorithms for scientific literature analysis.
17. **Infrastructure Development:** Advocating for more resources like GPU clusters and training in AI model deployment.
18. **AI in Sports Performance:** Using computer vision and AI for analyzing athletes' movements and injury patterns.
19. **Semantic Network Analysis:** Probing GLLMs to understand similarities between human and machine semantic networks.
20. **Educational Applications:** Integrating AI into classrooms to enhance teaching and learning experi-

ences.

#### B.3.2 Research Goals

What are your future research goals in AI over the next 1-5 years?

1. **AI in Classroom Assessment:** Exploring AI's potential to support assessment in educational settings.
2. **Instructional Planning with AI:** Utilizing AI for better instructional planning in education.
3. **AI-Enhanced Pedagogy:** Investigating how AI can improve teaching methods and pedagogy.
4. **AI in Healthcare Diagnostics:** Applying AI for diagnostics, particularly in conditions like cervical cancer.
5. **AI for Patient Outcome Predictions:** Using AI to predict patient outcomes in healthcare.
6. **Operational Improvements with AI:** Implementing AI for operational efficiency and performance enhancement.
7. **Automated Image Detection:** Employing AI for image detection in research, such as identifying black market goods.
8. **AI in Experimental and Computational Methods:** Linking AI with experimental and computational approaches for integrated research.
9. **AI for Emulating Models:** Utilizing AI to emulate complex models, especially in computationally intensive fields.
10. **AI as a Decision-Aide:** Exploring AI's role as an aid in decision-making contexts, like eyewitness identification.
11. **Ethical AI in Health Data Analysis:** Integrating ethical considerations and community input in health-related AI data analysis.
12. **AI as a Research Assistant:** Examining AI's potential to function as a research assistant.
13. **AI in Biomaterials Design:** Collaborating with AI experts to design healthcare-related biomaterials.
14. **Privacy Audits in AI:** Conducting meaningful privacy audits within AI systems.
15. **AI and Human Researchers:** Demonstrating how AI complements human researchers' contributions.
16. **AI in Brain Imaging:** Applying AI for designing and analyzing brain imaging data.
17. **Responsible AI:** Advancing AI that adheres to ethical standards and responsible usage.
18. **Physical Intelligence in Machines:** Developing physical intelligence capabilities in AI systems.
19. **Mathematics of AI Models:** Understanding the underlying mathematical principles of AI models.
20. **AI in Biology and Health:** Evaluating AI's application in biology and healthcare fields.
21. **AI in Literature Analysis:** Using AI to assist in analyzing literary works.
22. **AI for Neurodevelopmental Disorder Research:** Building AI models to identify disease signatures from biological datasets.
23. **AI in Qualitative Research:** Incorporating AI for theme analysis in qualitative research.
24. **AI for Cybersecurity:** Enhancing cybersecurity solutions through AI and machine learning.
25. **AI in Language Learning:** Leveraging AI to aid language learning and engagement with non-English languages.
26. **AI in Business and Media:** Focusing on AI at the intersection of business, media, entertainment, and the arts.

27. **AI's Costs in Human-Centered Work:** Investigating the impact of AI on human-centered work and its mediation in home relationships.
28. **AI in Civil and Environmental Engineering:** Applying AI in civil engineering, particularly for water resources management.
29. **AI for Data Analysis in Education:** Exploring AI's role in enhancing data analysis for educational research and classroom assessment.

## B.4 Resource Needs

Do you have the resources you need for the AI-based research you want to do?

1. **Sufficient Resources:** Several respondents – roughly half – believe they have the necessary resources for their AI research, with a number of affirmative responses such as "I believe so," "Yes," and similar expressions.
2. **Hardware Limitations:** Some researchers point out a lack of adequate hardware, specifically mentioning the need for better GPUs and more computational power to support core AI research.
3. **Computational Resource Concerns:** Issues with computational resources, such as limited access, competition for resources, and the need for more stable and robust high-performance computing facilities, are highlighted.
4. **Data Access and Storage:** Concerns about data access, storage limitations, and the reliability of current computing systems like Rivanna are mentioned, indicating a need for improved infrastructure.
5. **Cost of Resources:** The cost of accessing computing resources and the affordability of service units are concerns for some, affecting their ability to conduct AI research.
6. **Need for Institutional Support:** Researchers express a desire for institutional support, including access to AI tools like ChatGPT and assistance with IT tasks.
7. **Collaboration and Training:** There's a call for more opportunities for collaboration and training in AI, both within UVA and with external partners, to enhance research capabilities.
8. **Specific Resource Needs:** Some responses highlight very specific needs, such as access to more GPUs, improved subscription access to AI models, and enhanced collaboration tools.
9. **Ethical and Policy Concerns:** A few researchers are waiting for university policies on AI use in research and teaching before advancing their AI projects, indicating a need for guidance and ethical frameworks.
10. **External Dependencies:** Some researchers rely on external connections for computing resources, pointing to a potential gap in internal support.
11. **Diverse Research Needs:** The diversity of responses reflects the varied nature of AI research across disciplines, with different researchers requiring different types of support, from computational power to collaborative opportunities and ethical guidance.

## B.5 Additional Comments

### Topics

1. **Clear Policies for AI Research:** Need for clearer policies regarding AI research and its use in classrooms.

2. **Showcasing AI Research:** Suggestions to highlight ongoing AI research at UVA through bulletins or showcases.
3. **Centralized AI Resources:** Proposal for a centralized repository of GenAI information and tutorials.
4. **Investment in Existing Strengths:** Caution against solely chasing new trends without bolstering existing strengths.
5. **Leadership in AI Ethics:** Potential for UVA to lead in AI ethics and governance.
6. **Societal Consequences:** Emphasis on considering AI's societal impacts.
7. **AI in Healthcare:** Suggestions for AI applications in healthcare and patient outcomes.
8. **Institutional AI Access:** Addressing the need for institutional access to AI tools to avoid privacy concerns.
9. **AI Training for Faculty/Students:** Recommending training for UVA community on AI model usage.
10. **AI Research Visibility:** Suggestions to increase visibility and interdisciplinary collaboration in AI research.
11. **AI as a Research Tool:** Discussion on AI's role as a tool in research, particularly for non-AI specialists.
12. **Ethics in AI:** Highlighting the importance of ethics in AI research.
13. **AI and Teaching:** Exploring the link between AI research and its application in teaching.
14. **AI in Humanities:** Addressing the relevance of AI research to humanities and non-STEM fields.
15. **Critical and Anti-AI Research:** Advocating for research that critically examines or opposes AI.
16. **Generative AI:** Addressing the hype around generative AI and the need for balanced commitment.
17. **AI and Cultural Phenomena:** Considering AI as a technological, social, and cultural phenomenon.
18. **Cross-disciplinary AI Conversations:** Encouraging dialogue between AI users and those studying AI's broader impacts.
19. **Predictive Models and Ethics:** Emphasizing the need to couple predictive AI models with explainability and ethical considerations.
20. **AI Education for Researchers:** Highlighting the need for education on AI's effective and responsible use, especially for researchers without a computational background.

**Opinions**

1. **Need for Policy and Governance:** There's a call for clearer policies and governance structures concerning AI research and its application, particularly in educational settings. This includes a desire for UVA to take a leading role in the ethics of AI research.
2. **Promoting Interdisciplinary Collaboration:** Some respondents suggest that UVA could enhance AI research visibility and foster interdisciplinary collaboration by showcasing current AI research and facilitating connections between departments.
3. **Accessibility and Training:** There's a concern about providing easier access to AI tools for faculty and students, avoiding privacy and security issues, and offering training to ensure the UVA community can effectively leverage AI in their work.
4. **Balancing Innovation with Existing Strengths:** While recognizing the importance of AI research, some responses caution against neglecting existing strengths in pursuit of the latest AI trends, suggesting a balanced approach to innovation.
5. **Ethical Considerations:** Many responses emphasize the importance of ethical considerations in AI

research, suggesting that UVA has the potential to lead in this area and highlighting the need to consider AI's societal impacts.

6. **Critical Perspectives on AI:** Some voices advocate for a more critical approach to AI, suggesting that UVA should also support research that critically examines AI or is skeptical of its benefits, emphasizing the need to understand AI's potential downsides.
7. **Enhancing Education and Research:** There's a recognition of AI's role in enhancing education and research, with suggestions to integrate AI into teaching and to use AI as a research tool across various disciplines.
8. **Focus on Predictive Models and Explainability:** Some respondents highlight the importance of focusing on predictive models in AI and stress the need for these models to be complemented with efforts in explainability, interpretability, and ethics.

# C Bibliometric Analysis Methods

We use a combination of automatic text classification with Retrieval-Augmented generation and zero-shot text classification to develop topic scores for papers in the dataset. We used RAG to identify the most representative themes (or topics) across all publication records extracted via OpenAlex. Zero-shot classification was used with the goal of computing scores per each one of the topics to be further analyzed. The automatic text classification of themes was implemented using TinyLLAMA[25], a lightweight version of Meta's Llama-2 model with only 1.1B parameters. Zero-shot classification was implemented using the DestilRoBERTa model.[26]

# D AI Ethics Organizations

Many university organizations have been established to address the ethical impacts of AI. Among practitioners of human-centered AI, two centers stand out:

**Stanford University Institute for Human-Centered AI** (https://hai.stanford.edu/):

- Mission: To advance AI research, education, and policy to improve the human condition.
- Membership: Led by faculty from multiple departments.
- Activities: Research on AI technologies inspired by human intelligence; studying, forecasting and guiding the human and societal impact of AI; designing and creating AI applications that augment human capabilities. Education and outreach directed at students and leaders leading to regional and national discussions that lead to direct legislative impact. The institute delivers a Graduate fellowship program as well. See the Annual Report 2022.

**Berkeley University Center for Human-Compatible AI** (https://humancompatible.ai/):

- Mission: To develop the conceptual and technical wherewithal to reorient the general thrust of AI research towards provably beneficial systems.

[25]Peiyuan Zhang, Guangtao Zeng, Tianduo Wang, and Wei Lu. *TinyLlama: An open-source small language model*, arXiv:2401.02385, 2024.

[26]Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. *DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter*, arXiv:1910.01108, 2019.

- Membership: Core faculty include mainly Computer Science with some Psychology.
- Activities: Research on developing and communicating AI systems that defer to humans in light of that uncertainty; the foundations of rational agency and causality, value alignment and inverse reinforcement learning, human-robot cooperation, multi-agent perspectives and applications, and models of bounded or imperfect rationality; adversarial training and testing for ML systems, various AI capabilities, topics in cognitive science, ethics for AI and AI development robust inference and planning, security problems and solutions, and transparency and interpretability methods. Also publishes for general audiences and advises governments and international organizations, offering insight on a variety of individual-scale and societal-scale risks from AI, such as pertaining to autonomous weapons, the future of employment, and public health and safety. See a list of Publications.

Some notable AI ethics organizations include:

**Oxford Institute for Ethics in AI** (https://www.oxford-aiethics.ox.ac.uk/):

- Mission: To address "the ethical implications of AI from a philosophical and humanistic perspective" by bring together eminent thinker from a variety of fields and following the path laid by biomedical ethics.
- Membership: World-leading philosophers and other experts in the humanities with the technical developers and users of AI in academia, business and government. The Institute is part of thePhilosophy Faculty.
- Activities: Research broadly on democracy, governance, human tights, human well-being, the environment, and society.

**Emory Center for Ethics** (https://ethics.emory.edu):

- Mission: Advancing humanity with ethical AI. Engages with a wide range of stakeholders, from students to CEOs.
- Membership: Core faculty include experts in bioethics, general ethics, and pedagogy. These are joined by a broad range of senior faculty fellows.
- Activities: Offers an Online Certification Course in Ethical AI, holds AI Ethics brown-bag talks, and many other engagements. See this document for a full description of activities.

**The Artificial Intelligence Academic Initiative Center Center at the University of Florida (AI2)** (https://ai.ufl.edu/about/ai2-center/):

- Mission: AI2 is the university's focal point for academic initiatives related to AI and data science.
- Membership: A broad range of faculty from philosophy, business, law, engineering, and other disciplines.
- Activities: A working Group in AI Ethics and Policy; a required ethics course (*Ethics, Data, and Technology*) for all certificates and majors related to AI and data science; educational partnerships (including K-12), professional development. Faculty also pursuse research and publish on topics relating to ethical AI.

**Georgetown Center for Digital Ethics** (https://digitalethics.georgetown.edu/ai-artificial-intelligence-ethics):

- Mission: To help the world cope with technological challenges and opportunities, from artificial intelligence to cybersecurity, that are outstripping social, legal, and ethical guidelines.

- Membership: A university-wide community of scholars, including professors of philosophy, computer science, and law.
- Activities: A research area devoted to AI Ethics research focused on understanding the opportunities and risks of AI systems to employ them justly and legitimately in a democratic society. Sponsors an Annual *Penn-Georgetown Digital Ethics Workshop* featuring new research on digital ethics by rising early-career researchers from various disciplinary backgrounds including philosophy, business ethics, and computer science.

## E Policy Documents

The following documents and resources were identified during the interview with policy stewards as relevant to the topic of AI and research.

- RES-002: Ownership, Retention, Safeguarding, Management, and Transfer of Research Records: Researchers (e.g., faculty, research staff, fellows, assistants, technicians, students, and volunteers) shall maintain complete and verifiable records of the procedures they have followed in pursuing all research, and the subsequent data they have thereby obtained. Questions about this policy should be directed to the Office of the Vice President for Research.
- RES-009: Solicitation, Clearance, Acceptance and Ongoing Management of Sponsored Programs Section 3. Sponsored Program Contracts: Sponsored program contracts must be executed by an authorized signatory in OSP (see University policy FIN-036: Signatory Authority for Executing University Contracts). Sponsored program contracts include, but are not limited to, certain research awards (i.e., contracts and collaborative research and development agreements), collaboration agreements, memoranda of agreement or understanding, nondisclosure agreements, and agreements related to the transfer and use of materials and/or data when associated with the conduct of research. Questions about this policy should be directed to the Office of Sponsored Programs.
- IRM-003: Data Protection of University Information: Applies to data in any format. Users must comply with all University policies, standards, and procedures for the data to which they have been granted the ability to view, copy, generate, transmit, store, download, or otherwise acquire, access, remove, or destroy. Users must also meet any additional compliance requirements for data protection stipulated by various governmental, legal, or contractual entities, including, but not limited to, those defined for classified information, Controlled Unclassified Information (CUI), International Traffic in Arms Regulations (ITAR) covered data, Payment Card Industry (PCI) regulated data, Health Insurance Portability and Accountability Act (HIPAA) covered data, and Federal Educational Rights and Privacy Act (FERPA) covered data. Questions about this policy should be directed to University Information Security (InfoSec).

*Other relevant documents*

- RES-007: Human Research Protection ProgramSection 5. Research Data Security Section 5. Research Data Security
- RES-011: Investigator Eligibility Requirements and Responsibilities Related to Sponsored Programs
- FIN-017: Consistent Treatment of Sponsored Program Costs

- FIN-043: Managing Export and Sanction Compliance in Support of University Activities Section 6. Research Data Security
- IRM-004: Information Security of University Technology Resources
- PROV-032: Visiting Graduate ResearchersSection H. Data Management
- PROV-008: Teaching Courses for Academic Credit Section 11
- HRM-025: Professional Service and External Consulting for University Staff EmployeesSection 5. Use of Highly Sensitive Data and Intellectual Property
- HRM-033: Employment of Professional Research Staff Section 2.c
- FIN-021: Internal Control responsibility for safeguarding data
- IRM-012: Privacy and Confidentiality of University Information
- PROV-016: Out-of-State Educational Activities

*Resources*

- Office of Sponsored Programs.
- Office of Export Controls.
- University Information Security (InfoSec)
- University Records Management Office
- Medical Center Risk Management
- Office of Institutional Research and Analytics.

## F Conversation with a Friendly IRB Bot

The following is an unedited transcript of a conversation between a human used and an AI bot that was created to answer questions about the IRB. We have replicated as closely as possible the formatting used by the bot, but within the constraints of this document. Users with a Poe account, can try out the bot themselves at https://poe.com/UVAIRBDemoBot (but, of course, as noted in its welcome message, it is only intended as a demonstration of the readily-available capabilities of generative AI—please do not rely on its answers).

The bot was provided with all of the content from the VPR's IRB website (https://research.virginia.edu/irb-sbs), just by crawling all pages linked from the starting URL. The responses include links to those webpages, shown as [3] (due to the use of the crawled data, the actual links cannot be replicated here, but seem to point to relevant sources).

The underlying language model is OpenAI's GPT-4 (although our experience is that Anthropic's Claude-3 performs competitively with this model). The instruction prompt is simple, "You are a very helpful assistant. Do not make up answers. If the answer is unclear from the available sources, encourage the questioner to ask the IRB. Be friendly.".

**UVAIRBDemoBot:**

Hello - this is the IRB chatbot! This is just for demonstration purposes only, please do not take any of these answers as definitive or interpret them as legal advice - I'm just a stupid chatbot using a probablistic model to generate text responses.

**Human User:**

what are the steps to getting an IRB approved for human subject experiments?

**UVAIRBDemoBot:**

To get an IRB approval for human subject experiments, you should follow these steps:

1. **Write Your Protocol for IRB Approval**: Begin by writing a detailed study protocol that outlines your research plan, including objectives, methodology, data collection tools, and participant recruitment and consent processes [1].
2. **Complete Required Trainings**: All researchers involved in the study must complete CITI Training, which is valid for three years. This training is necessary to ensure that researchers understand the ethical considerations and regulations surrounding human subjects research [3].
3. **Create and Submit a New iProtocol**: Use the iProtocol system to create and submit your study for IRB review. Make sure to review the iProtocol Question Guide for assistance and upload all relevant documents, such as data collection tools and consent forms [1][4].
4. **Obtain Dean's Approval**: After working with an IRB pre-reviewer on your protocol, you will need to secure the Dean's approval. This step is completed within the iProtocol system under the section "Permission to Access Data Source and Participant Group" [1].
5. **Understand Review Processes and Vocabulary**: Familiarize yourself with the IRB review processes, including the difference between expedited and full board reviews, and the specific vocabulary you will encounter during the review [2].
6. **Submit to IRB for Review**: Once your protocol is in its final stage, with all necessary approvals and training completed, submit it to the IRB for review. If your study involves more than minimal risk, it will need to be reviewed at a Full Board Meeting, which are held twice monthly [1][5].
7. **Address Any Feedback**: The IRB may provide feedback or request changes to your protocol. Address these promptly and resubmit your revised protocol for approval.

Remember, the IRB's role is to ensure that your study complies with ethical standards and federal regulations for the protection of human subjects. It's important to plan for enough time to complete the IRB review process when designing your study. If you have any specific questions or concerns, you should contact the IRB office directly for assistance [2][3].

**Human User:**

I need to do experiments with prisoners. Are there any special steps I should take?

**UVAIRBDemoBot:**

When conducting experiments with prisoners, there are indeed special steps and considerations you must take due to the vulnerable status of this population:

1. **Understand Federal Regulations**: Familiarize yourself with Subpart C of the federal regulations, which provide specific instructions on how prisoners can be included in research studies [5].
2. **Ensure Appropriate IRB Composition**: When submitting your protocol for review, ensure that the IRB has a prisoner representative or someone with the appropriate background and experience to review prisoner studies. This person may not be a regular member of the board and serves on a needs basis [1].
3. **Avoid Undue Influence**: Payments and benefits of the study must not unduly influence prisoners to participate. Consider the impact of any benefits in the context of the prison environment [1][3].
4. **Fair and Impartial Selection**: Recruitment and selection of participants must be fair and impartial, and the consent process must be voluntary and free from coercion by prison authorities [2][3].
5. **Conduct Recruitment and Consent**: Recruitment and consent must be conducted by the researcher, not by prison staff, to avoid coercion. Provide a recruitment plan and consent procedure that encourages voluntary participation [2][3].
6. **Communicate with Prison Authorities**: Contact the prison before submitting your protocol to discuss any rules or regulations regarding conducting research and any potential hurdles to the consent process [2].
7. **Use Understandable Language**: The consent form and consent session must be presented in language that is understandable to the participants [2].
8. **Assess Risks**: Ensure that the risks of the research are comparable to risks that would be acceptable to non-prisoner participants [3].
9. **Research Categories**: Prisoners can only participate in studies that are about the causes, effects, and processes of incarceration, studies of the prison institution or the incarcerated person, or research on conditions particularly affecting prisoners as a class [4][5].
10. **Post-Study Care**: Ensure that any need for post-study care for prisoners is appropriately provided, taking into account the varying lengths of prisoners' sentences [6].
11. **Modify Protocol if Necessary**: If a participant becomes a prisoner during

the study, you must submit a modification describing the situation for IRB review [4].

12. **Contact IRB Immediately**: If an incident occurs while conducting your study that affects the status of participants as prisoners, contact the IRB office immediately for assistance [1].

It's crucial to give the IRB adequate time to schedule the review of a prisoner study and to ensure that all these special considerations are addressed in your protocol submission [1]. If you have any questions or need further guidance, you should contact the IRB directly.

**Human User:**

That sounds like a lot of trouble. Instead, I'll just use an AI chatbot.

**UVAIRBDemoBot:**

Using an AI chatbot for your research could indeed be a simpler alternative, as it avoids the complex ethical and regulatory challenges associated with conducting experiments with prisoners. When using AI technology like a chatbot, you don't have to navigate the stringent requirements for informed consent, fair participant selection, and the potential for coercion that come with prisoner research [1][2][3]. Additionally, you won't need to coordinate with prison authorities or adapt your study to the specific and often restrictive prison environment [2][3].

However, it's important to consider the objectives of your research and whether an AI chatbot can meet those needs. If your study aims to understand human behavior, particularly that of prisoners, an AI chatbot may not provide the same depth of insight as interacting with real human participants. On the other hand, if your research can be conducted without the need for human subjects, using AI could streamline your process and avoid the ethical complexities altogether.

Remember to ensure that your use of AI and any data collected complies with relevant privacy and ethical guidelines, as these considerations are important in all types of research. If you have any doubts or questions about the suitability of an AI chatbot for your study, consulting with the IRB could provide valuable guidance.