AI-VERDE : A Gateway for Egalitarian Access to Large Language Model-Based Resources For Educational Institutions

Paul Mithun, **Enrique Noriega-Atala**, **Nirav Merchant**, and **Edwin Skidmore**

Data Science Institute
University of Arizona
{mithunpaul,enoriega,nirav}@arizona.edu

Abstract

We present AI-VERDE , a unified LLM-as-a-platform service designed to facilitate seamless integration of commercial, cloud-hosted, and on-premise open LLMs in academic settings. AI-VERDE streamlines access management for instructional and research groups by providing features such as robust access control, privacy-preserving mechanisms, native Retrieval-Augmented Generation (RAG) support, budget management for third-party LLM services, through both conversational web interface and API access, etc. In a pilot deployment at a large public university, the University of Arizona AI-VERDE demonstrated significant engagement in various educational and research groups, enabling activities that would typically require substantial budgets for commercial LLM services with limited user and team management capabilities. To the best of our knowledge, AI-Verde is the first platform to address both academic and research needs for LLMs within an higher education institutional framework. Further, several of the solutions provided by AI-VERDE , albeit for educational institution perspective, can be easily extended to any closed community of users like in a corporate environment.

1 Introduction

Large Language Models (LLMs) like ChatGPT (OpenAI, 2022), GPT-4 (OpenAI, 2023), Mistral (Jiang et al., 2023a), Llama (Dubey et al., 2024), etc., have rapidly emerged as transformative tools that demonstrate significant capabilities across a broad spectrum of applications, including natural language processing, content generation, and educational support. Their powerful capabilities have captivated the attention of users and has led to a huge creative exploration and novel applications that span most disciplines.

Due to this accelerated advance and acceptance of the LLM technology by the hoi polloi, Universities and colleges are also currently under pressure to integrate this into academic settings. However, integrating LLM technology into academic settings, especially higher education institutions, faces several unique challenges, such as, the minimum technological know-how expected from the users, privacy concerns, limited access to specialized knowledge, intellectual property rights etc. For example, a Professor aiming to incorporate commercial LLMs like ChatGPT into their coursework often encounters high costs, as well as difficulties in managing usage across students and generating access tokens. Even for the free open-source LLMs like Llama, the required supporting frameworks such as knowledge of a programming language pose a significant barrier, especially Faculty coming from a non-STEM background. Even with basic programming knowledge, using an LLM effectively for advanced courses requires fine-tuning or setting up a Retrieval Augmented Generation (RAG) (Lewis et al., 2020) pipeline, both of which involve complex software and hardware requirements. Moreover, issues related to intellectual property, such as the potential public dissemination of copyrighted textbooks, add further challenges to seamless integration. The list goes on.

In this context these are the unique contributions of our work:

- We present a detailed survey capturing several problems that prevent smooth adoption of LLM based technology at Universities and other higher education institutions

- We present AI-VERDE [1], an LLM-platform-as-a-service (LLMPaaS) which addresses most of these concerns. A pilot version of the same is already being incorporated by several Professors and Researchers at the University of Arizona.

[1]Video demo: https://youtu.be/hPTiZlcUiZo

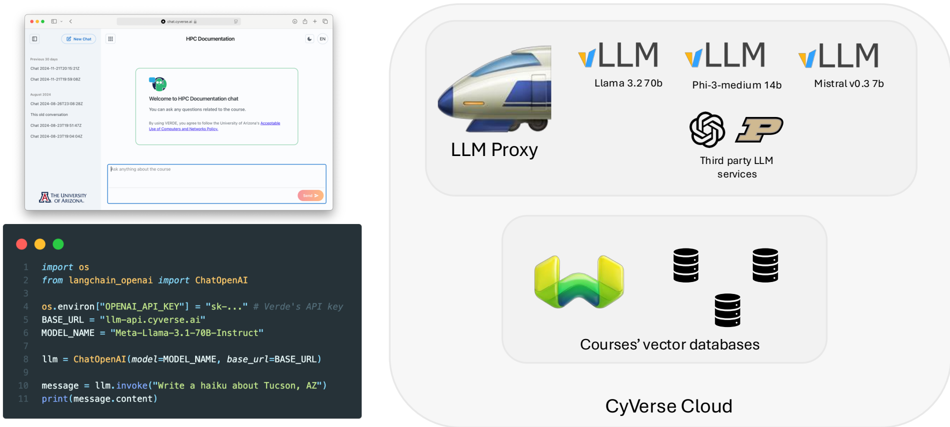

Figure 1: Architecture Diagram of AI-VERDE . The left hand side of the diagram represents the *frontend*, which consists of the conversational web interface, depicted at the top, and a snippet of code with an example of how to programmatically connect to AI-VERDE using industry-standard python software packages. The right side depicts the *backend* elements, and illustrates multiple different models running with vLLM, as well as a proxy to commercial models, all exposed to clients through LiteLLM. The backend also contains our managed instance of the Weaviate vector database manager, which houses the different vector databases, corresponding to each course, enabled in AI-VERDE .

- We do a comparative study with other commercial options and show that AI Verde provides a much lower cost egalitarian gateway to AI tools, and can thus democratize the access to LLMs in colleges and University campuses.

To the best of our knowledge this is the first that such a comprehensive investigation, and a product catering to higher education institutions to holistically address these limitations, as an integrated platform, is being done.

The rest of this work is organized as follows. First in section 2 we discuss some prior art related to LLMs, specifically its usage in academic settings. Then in section 3, we will discuss the logistical details of how the survey was conducted. Then in section 4 we highlight some of the major concerns raised in this survey. In section 5 we introduce AI-VERDE and details of its internal architecture. Following that, in section 6, we show how AI-VERDE addresses several of these problems that impede the easy installation and adaptation of LLMs /AI infrastructure on University campuses. This is followed by discussion, future work and conclusion sections.

2 Related Work

LLMs have proliferated quickly since the inception of the transformers architecture (Vaswani et al., 2017). As the size of these models scaled very quickly (Simon, 2021), the increased computing requirements have driven users towards deploying modern LLMs in client-server environments. As a result, multiple open source software projects designed to efficiently serve models have emerged[2][3][4]. These software systems model their application programming interfaces (API) after OpenAI's proprietary API, which allows increases the interoperability and modularity in software development. LLM gateway software such as LiteLLM leverages this interoperability to expose a unified API proxy service that manages access and usage to multiple models. LLMs are very often used in tandem with external information repositories and databases to build *retrieval augmented generation* (RAG) applications (Lewis et al., 2020). Often, the information is stored as documents, which are encoded using transformer models designed to compute pairwise semantic similarity (Reimers and Gurevych, 2019). For this particular use case, multiple vector database management solutions exist that leverage optimized algorithms to compute exact or approximate nearest-neighbors search (Douze et al., 2024).

Platforms an models specifically to adoption of LLM in University settings are however limited. After the recent AI boom, triggered and accelerated by introduction of ChatGPT (OpenAI, 2022), several universities, government agencies

[2] https://ollama.com
[3] https://huggingface.co/docs/text-generation-inference/en/index
[4] https://github.com/vllm-project/vllm

and private institutions have established collaborations and consortia to reduce the barrier of entry to LLMs for learning groups. Arizona State University (ASU) recently partnered with OpenAI[5] to allow groups to apply for access to ChatGPT for research purposes. Purdue University's Anvil (Song et al., 2022) provides API access to open-source LLMs through NSF's ACCESS.

3 Survey details

To understand the pulse of the Faculty, students and staff in a university campus towards adoption of AI for education and research purposes, a survey was conducted in Spring 2024, at the University of Arizona.

To start with, a campus wide email was propagated first asking for participants who might be interested in giving feedback on AI adoption on campus. From these responses, 3 major persona groups were separated, Faculty, students and staff. From these, various modalities of information extraction were done, as detailed below.

It must be noted that this survey deviated from the standard modality of a single survey comprising the same questions, that would be filled by all users. Instead it was done in several different modalities based on various factors including count and convenience of the intended group. For example with Faculty, it was easier to do comprehensive hour long one on one interviews. While with undergrad students, since the real estate to cover was larger, focus groups were conducted, where the students were asked to debate about this topic.

In case of undergraduate students who responded showing interest a subsample of 53 students were picked up. This sampling ensured that it included a fair representation of students from different levels and modes of undergraduate journey, including, freshmen through seniors, students with multiple majors, spread across several different colleges etc. This ensured that the final sample of students picked represented a fair representation of every undergraduate persona group distributed across campus. The final 53 students were further divided into five groups. Each of these groups were assembled in person, and a group discussion was conducted, each an hour long. At the end of the hour they were asked to come up with their top 10 concerns regarding adoption of AI on college campuses. .

Meanwhile with the graduate students, since the sample size was lesser, eight in depth one- on-one interviews were conducted. Further, 27 others responded through email to some standard survey questions about AI needs and interests. Like the undergraduate sample, it was ensured that this represented graduate students across various colleges and disciplines, at various levels of their journey, including masters and Phd students.

Amongst the Faculty and staff of University of Arizona, 41 one-on-one interviews were conducted to understand in-depth their needs and wish list with AI tools, in supporting their research and instructional needs. Do note that the starting sample was less than 5, but these snowballed into referrals to other Faculty who they thought would be definitely interested in responding to these. These one-on-one interviews included 12 research Faculty and 2 UA library staff who were already using AI tools.

Apart from all these efforts, an email survey with open-ended questions was sent out to 112 Faculty, staff, and graduate students who were part of the Faculty Learning Community (FLC)[6] at University of Arizona. From this group we received 57 responses. Further there was an open call for AI needs, ideas and scenarios that went out to nice groups at the University of Arizona like the AI^2 Task Force[7], University Center for Assessment, Teaching and Technology [8], and to the staff of University of Arizona Libraries [9], and some to non-AI focused FLCs like Agriculture, Life and Veterinary Sciences, and Cooperative Extension[10] and CALES [11]. This yielded 194 responses. So in total, 372 participants from various walks of life in our University campus were interviewed for this survey. In addition, document analysis and prior art search was done to identify potential uses and the current state of private or public LLMs in Higher Education.

[5] https://www.insidehighered.com/news/tech-innovation/artificial-intelligence/2024/05/21/unpacking-asus-openai-partnership-and

[6] FLC is a peer-led group of Faculty members (typically 6 to 12) who engage in an active, collaborative, year–long program, structured to provide encouragement, support, and reflection in teaching and learning.

[7] https://artificialintelligence.arizona.edu/about-us

[8] https://ucatt.arizona.edu/

[9] https://library.arizona.edu/

[10] https://alvsce.arizona.edu/

[11] https://research.cales.arizona.edu/

4 The Problems

The most prominent concerns identified through this survey, emphasizing the practical and ethical issues that arise when deploying such AI platforms in University campuses, were as follows.

4.1 Intellectual Property, Privacy and Content Ownership

The primary concern that was raised in the survey, agreed unanimously by all Faculty, staff, and students, was the lack of privacy and control over their data. This is because the current typical providers of LLM, especially the commercial platforms, often store user data, such as copyrighted course materials or sensitive queries, on corporate servers, where it may be used for downstream model training (Yao et al., 2024). This raises issues of data security, intellectual property rights, and confidentiality, particularly in educational settings.

4.2 Limited Access to Specialized Knowledge

General-purpose LLMs platforms, while effective for general information retrieval, often fail to meet domain-specific needs (Minaee et al., 2024). Specialized chatbots for fields like medicine or law are typically limited to answering frequently asked questions and lack the conversational depth required for complex academic queries (Yigci et al., 2024). These limitations reduce accuracy and trust, diminishing their utility in research and education.

4.3 Authorization and Authentication

Another major challenge with commercial chatbots in academic settings is their inability to integrate smoothly with existing university platforms, such as authentication and authorization protocols of the respective learning management systems (LMS) (Oliveira et al., 2016). In the Universities in the United States, for instance, typically proprietary content like grades or course materials in an LMS often requires authentication via dedicated platforms ensuring access is restricted to individuals with university-affiliated email addresses. Furthermore, post-authentication, there are additional authorization requirements, such as ensuring only students registered for a course can access its related chatbot.

4.4 Equity and Resource Constraints

The freemium[12] business model that several AI platforms have adopted, often disadvantage financially constrained students and institutions by restricting advanced features to paid subscriptions. Token-based pricing and manual on-boarding processes, even in NSF-funded AnvilGPT[13], hinder scalability, making it difficult for educators to manage large cohorts. These challenges underscore the need for automated enrollment and management solutions in educational LLM applications while addressing financial compliance concerns in cloud-based resources. Additionally, the increasing reliance on cloud-based resources has introduced significant concerns over budget management and financial compliance for academic departments.

4.5 Anonymity And Confidentiality Of Sensitive Data

Compliance with regulatory standards such as Health Insurance Portability and Accountability Act (Act, 1996) (HIPAA), Family Educational Rights and Privacy Act (Rights and Act, 2014) (FERPA), and Data Use Agreements (DUAs) in research and education contexts often prohibits the use of external services for handling sensitive data. In several cases respective institutions have to ensure that external service providers do not inadvertently compromise data privacy or violate agreements of the end-users they cater to. For example, many academic research projects involve human subjects, requiring compliance with strict ethical and legal standards like IRB approval to ensure data privacy and confidentiality. However, the usage of commercial large language model (LLM) platforms raises significant concerns due to ambiguous data handling practices (Yao et al., 2024; Wang et al., 2023b; Jaff et al., 2024), risking the exposure of sensitive information, including, possibly, Personal Health Information (PHI). Additionally, limiting the tracking of individual user activities is also a priority in certain settings and projects.

4.6 Access to dedicated hardware

Effectively harnessing the power of LLMs requires fine-tuning on specialized datasets or integrating them with Retrieval-Augmented Generation (RAG) systems. However, setting up a RAG system with

[12] freemium is a type of business model that offers basic features of a product or service to users at no cost and charges a premium for supplemental or advanced features.

[13] https://anvilgpt.rcac.purdue.edu/

Feature	*ChatGPT*	*Gemini*	*Anvil*	*AI-VERDE*
Privacy-Preservation	No	No	No	Yes
Built-in guardrails	Yes	Yes	Unknown	Yes
On-Premises Deployment	No	No	No	Yes
Native RAG support	Limited	Limited	No	Yes
Instructional groups management	No	No	No	Yes
Hosted models customization	No	No	No	Yes

Table 1: Feature comparison between AI-VERDE and other LLM comparable platforms. ChatGPT and Gemini are representative of commercial offerings, while Anvil is representative of academic offerings.

large-scale LLMs like LLaMA-3 or GPT-4 demands significant hardware resources, including high-end GPUs (e.g., NVIDIA A100s or H100s), substantial VRAM (48–150GB), and system memory (128GB or more). Additionally, fast storage (NVMe SSDs) and potentially distributed architectures are necessary for large-scale deployments, making such setups costly, with budgets often exceeding $50,000—a prohibitive expense for most campus researchers.

4.7 Steep Learning Curve

While AI and LLM might not be that intimidating to the STEM majors, for a student or Faculty coming from disciplines like humanities and social sciences even python programming will be intimidating, let alone details of AI. This steep learning curve presents a significant entry barrier to smooth usage of AI technology on campus. Specifically, the primary uses of LLMs on university campuses are supporting teaching and enhancing research. While Faculty and researchers are the main users, they often face challenges due to a lack of programming or AI expertise, particularly those from non-STEM disciplines. Current LLM solutions require significant technical skills for fine-tuning on specialized data or implementing RAG, making them inaccessible to many.

4.8 Hallucination and Misinformation

Another key concern highlighted in the survey is the issue of hallucination, where general-purpose chatbots produce responses that are plausible but factually incorrect. Research has shown that this happens when chatbots struggle to find answers to questions. This issue is particularly problematic in academic settings, as students may unknowingly rely on and learn from incorrect information, leading to negative learning outcomes. The issue is further complicated by broader ethical concerns such as enabling cheating, plagiarism, and breaches of academic integrity standards.

4.9 Guard rails

Another major concern raised in the survey was the possibility of inappropriate language used, either by the user or by the LLM in its response.

4.10 Prompt Engineering

Effective utilization of LLMs requires careful prompt engineering (Zhou et al., 2022; Wang et al., 2023a). Note that prompt engineering doesn't need programming or AI knowledge, so can be done by a Professor or Researcher with minimal training. However, prompt engineering gets trickier within a RAG system. That is because the prompt supplied by the user on a typical chatbot based interface is used to query relevant documents from the vector database. However, the actual prompt presented to the LLM is different and the end-user doesn't have access to it. This lack of access to the final prompt restricts the end user.

4.11 On-boarding And Enrolling

As shown above, while the challenges posed by technical issues itself are high, another key challenge in provisioning LLM based access for users in any systematic structured teaching environment (e.g.students taking courses, workshops etc at a University) is the logistical challenge. Specifically, on-boarding or enrolling participants at the beginning of a term (e.g. semester) and often deprovisioning them when the event concludes. However, in the other currently available unified LLM-as-a-platform service solutions, this process is being done manually. For example, in the NSF funded Purdue University's Anvil (Song et al., 2022) each user has to create their individual account. Further even after registration, the approval and actual on-boarding typically takes twenty four to fourty eight hours. The survey showed that ensuring that this on-boarding process is done systematically for a class of say 50 students is a huge burden and hurdle for the Faculty, even if the resource (LLM-as-a-platform service) is provided at no cost.

5 AI-VERDE

The primary goal in designing of AI-VERDE was to create an egalitarian gateway for members of academia, to smoothly access all facets of AI technology. Hence, while at the core AI-VERDE comprises of a standard RAG pipeline to LLMs, Each individual sub-modules and component were made available to be accessed independently as micro-services, in a plug-and-play format.

Specifically, AI-VERDE [14] is a platform designed with the goal of providing seamless access to LLMs (especially RAG if need be) to the academic community through various means like chat interfaces, API etc. To achieve this functional requirement, we tapped into open source technologies to serve LLMs as well as developed custom software to allow the provisioning of a multi-tenant environment to accommodate diverse educational and research groups. In this section we detail several of the major architecture components of AI-VERDE which is also shown in Figure 1. The individual pieces are deployed and orchestrated with Kubernetes (Beda et al., 2014).

5.1 Backend: LLM serving

At the heart of AI-VERDE lies the LLMs to be exposed to our community. We use vLLM (Kwon et al., 2023) to persistently load LLM in a GPU cluster. vLLM allows us to serve open source models, such as Llama 3.2 (Touvron et al., 2023), Mistral (Jiang et al., 2023b) and Phi-3 (Abdin et al., 2024) and provide an API interface. Additionally, we leverage the advanced capabilities offered such as support for numerous model weight formats, integration with the HuggingFace Hub[15], support for LoRA adapters (Hu et al., 2022), and customized paged attention for increased throughput. Each LLM served through AI-VERDE is associated with a running instance of vLLM.

We use LiteLLM[16] as an *LLM proxy* to provide a unified, managed API access to all the individual vLLM instances. Intuitively, LiteLLM behaves as a *reverse proxy*: It exposes an OpenAI-compliant API access point that routes the requests to the corresponding LLM by its name. LiteLLM enables user access control through and usage metering through the use of API Keys. Additionally, we can also use the LLM proxy functionality to seamlessly meter access commercial LLM API providers such as OpenAI and Anthropic, or other research LLM services such as AnvilGPT[17]. We issue surrogate API keys to allow us to provide a fine-grained managed access and control budgets. To support Retrieval Augmented generation configurations, we host a Weaviate[18] vector database environment. Access to Weaviate is controlled too through the use of API keys.

5.2 Front-end: User facing web interface

We introduce a software interface designed to complement our model serving and storage features. This has two major components.

5.2.1 Conversational user interface

We provide a conversational user interface to enable quick access to an LLM, if that is deemed to be the requirement. The LLM can be configured either in *pass through* or *RAG* mode. The former passes the conversation directly to the LLM configured for the course's interface whereas the latter executes a RAG workflow using the vector database assigned to the course. The conversational interface manages the conversation's history and persists and retrieves past conversations in a similar fashion to comparable products such as ChatGPT.

5.2.2 Programmatic access

As an alternative to the conversational user interface, users are able to programmatically access the LLMs using the API keys provided via the web interface. Programmatic access is enabled thanks to the LiteLLM component and exposes an OpenAI-compliant API interface, which allows the use of popular third-party integration libraries to build AI-powered applications, such as Langchain[19] and Llamaindex[20].

5.3 Access control

At the user interface, the federated login and authentication is achieved through CILogon[21].

[14] https://chat.cyverse.ai/home
[15] https://huggingface.co/models
[16] https://github.com/BerriAI/litellm
[17] https://www.rcac.purdue.edu/news/6826
[18] https://github.com/weaviate/weaviate
[19] https://www.langchain.com
[20] https://www.llamaindex.ai
[21] https://www.cilogon.org/

Specifically, each individual user group typical has unique requirements, such as access to different models or access to a vector database. AI-VERDE caters to instructional and research groups and customizes a subset of LLMs and vector databases accessible to each. We implement the abstraction of groups as *courses*. Each course consists of a list of students and instructors. All users have access to their individual API key. Instructors also have access to the list of students and to the budget information of the course.

5.4 Document Intake

To support RAG workflows, we developed a document intake service that reads various file formats like MSWord, MSPowerPoint slides, PDF files etc., which are then used to generate a corresponding vector database persisted into AI-VERDE's Weaviate service. The document intake service runs independently of the backend and frontend components. Once a vector database is provisioned, it can be configured as part of a course to enable RAG in the conversational UI.

6 How AI-Verde Solves The Problems

This section outlines and details how AI-VERDE tackles several of the problems that impede the easy installation and adaptation of LLMs /AI infrastructure on University campuses. Table 2 provides a comparative visual overview of the features mentioned below, to some of the popular commercial and open source alternatives.

6.1 Intellectual Property, Privacy and Content Ownership

The primary concern that was raised in the survey, agreed unanimously by all Faculty, staff, and students, was the lack of privacy and control over their data. AI-VERDE addresses these concerns by ensuring all data, including user queries and uploaded content, is processed entirely on-premises within a secure infrastructure. Unlike commercial platforms, AI-VERDE does not store or reuse queries for model training and disables personalization features by default to prioritize privacy. These features are only activated with explicit user consent, maintaining strict control over sensitive academic and personal information.

6.2 Limited Access to Specialized Knowledge

AI-VERDE overcomes this challenge by leveraging RAG, an AI methodology that retrieves contextually relevant documents based on user queries and primes the LLM with this information. By indexing specialized knowledge datasets, such as niche research papers or course materials, AI-VERDE enables accurate, context-aware responses without storing or training on user-provided data. This ensures data security while empowering researchers and educators to access precise, domain-specific insights.

6.3 Authorization and Authentication

As mentioned earlier, publicly available commercial and open-source platforms typically fail to meet the needs in a University, especially operating in isolation and lacking flexibility to align with specific university protocols. AI-VERDE addresses these concerns by supporting seamless integration with university authentication platforms through CILogon, allowing registered users of University LMS to access services without added complexity. Additionally, AI-VERDE simplifies inter-university collaboration by offering a plug-and-play authentication interface, enabling smooth integration and secure cross-institutional collaboration.

6.4 Equity and Resource Constraints

Equity challenges raised are solved by AI-VERDE by providing free access to itself, for anyone on a given University campus. This is achieved by using open-source components, RAG technology, and partnerships with cost-effective hardware providers like CyVerse (Swetnam et al., 2024) and NSF's Jetstream2 (Hancock et al., 2021), ensuring minimal operational costs. It also automates budget management by enabling Faculty to allocate class-specific funds and distribute API keys efficiently. Moreover, AI-VERDE reduces the administrative burden by handling routine queries with chatbots, freeing Faculty to focus on research, teaching, and impactful academic work, thus providing a literal egalitarian gateway[22].

[22] It's egalitarian since any person on a University campus has equal access rights to AIVERDE. It is a gateway, since users can access any of their favorite state-of-the-art LLM, picked from the current extremely fast moving LLM production pipeline, while using the same access methodology.

Feature	*ChatGPT*	*Gemini*	*Anvil*	*AI-VERDE*
Privacy-Preservation	No	No	No	Yes
Built-in guardrails	Yes	Yes	Unknown	Yes
On-Premises Deployment	No	No	No	Yes
Native RAG support	Limited	Limited	No	Yes
Instructional groups management	No	No	No	Yes
Hosted models customization	No	No	No	Yes

Table 2: Feature comparison between AI-VERDE and other LLM comparable platforms. ChatGPT and Gemini are representative of commercial offerings, while Anvil is representative of academic offerings.

6.5 Anonymity And Confidentiality Of Sensitive Data

AI-VERDE addresses privacy concerns by operating entirely within a secure, on-premises infrastructure, eliminating risks of external data transfer. It integrates with Soteria[23], a secure data analysis enclave for HIPAA compliance, ensuring secure data processing. Additionally, its gateway model abstracts user interactions, preventing commercial providers from tracking or collecting individual user data, strengthening privacy and supporting ethical data management practices.

6.6 Access to dedicated hardware

AI-VERDE alleviates the prohibitive expense for most campus researchers and faculty by providing a fully equipped hardware infrastructure with pre-loaded LLMs. Further it features elastic hardware allocation tailored to user needs and integrates seamlessly with NSF Jetstream, CyVerse, and cloud services like AWS [24] and Azure [25], enabling efficient scaling for diverse research and teaching workloads.

6.7 Steep Learning Curve

While AI and LLM might not be that intimidating to the STEM majors, for a student or Faculty coming from disciplines like humanities and social sciences even python programming will be intimidating, let alone details of AI. Note that not everything in AI-VERDE needs programming knowledge- everything is based on the need of the end user. AI-VERDE team understands this, and would like to meet the person where they are. Specifically, AI-VERDE addresses the steep learning curve barrier by offering a dedicated education and support team. The education support, includes workshops for users at all skill levels and personalized planning sessions with AI specialists. This helps non-technical users adopt AI while enabling advanced users to leverage APIs and microservices for tasks like fine-tuning and RAG. Through these initiatives, AI-VERDE democratizes AI use, reducing the steep learning curve associated with LLM adoption.

Further, recognizing that AI adoption varies across user needs, AI-VERDE provides personalized planning sessions with AI specialist consultants. These consultants assess individual requirements—whether for researchers using personal GPUs or educators integrating AI into teaching—and design tailored solutions to optimize resource use. They also guide users through integration and provide hands-on training, significantly lowering barriers to AI adoption.

Also for experienced users with programming knowledge, AI-VERDE offers direct API access and microservices, enabling advanced tasks such as fine-tuning and RAG. Additionally, workshops and training sessions empower technical users to maximize the platform's capabilities. For example some users (e.g researchers on campus) would like to use their Laptop or their own GPU and not use a high performance computing or Cyverse. So to support it, AI Verde meets the person where they are. Which is why the very first step in AI Verde will be a conversation with our AI Specialist consultants who will give you a plan and path on how best you can use the available resources.

6.8 Hallucination and Misinformation

Another key concern highlighted in the survey is the issue of hallucination, where general-purpose chatbots produce responses that are plausible but factually incorrect.

While not completely eliminating hallucination, AI-VERDE reduces it to a great extend by utiliz-

[23] https://soteria.arizona.edu/
[24] https://aws.amazon.com/
[25] https://azure.microsoft.com/

ing a combination of advanced prompt engineering techniques and forced refusal to reply. Specifically, through extensive trials, a prompt was developed to ensure the platform responds with a clear negative acknowledgment (e.g.,"That question is beyond my purview of current knowledge") when it cannot provide a correct answer, as opposed to hallucination. While we don't claim that the problem of hallucination has been completely solved, when combined with appropriate temperature settings and RAG, AI-VERDE grounds its responses in verified, retrieved documents. This approach ensures that the generated outputs are accurate, reliable, and aligned with the original source material with minimal hallucination .

6.9 Guard rails

Another major concern raised in the survey was the possibilty of inappropriate language used, either by the user or by the LLM in its response. To address this AI-VERDE incorporates robust solutions like Llama guardrails [26] to prevent biased or inappropriate outputs.

6.10 Prompt Engineering

Effective utilization of LLMs requires careful prompt engineering (Zhou et al., 2022; Wang et al., 2023a). But most of the typical providers of RAG system don't provide access to the final prompt sent to the LLM. To overcome this,AI-VERDE provides API and Chatbot based access to most of the state-of-the-art LLMs (both the paid commercial ones and the free open source ones), builds a RAG pipeline over a specialized vector database created from specialized knowledge provided by an end user. In fact chatbot is only one of the services provided by the platform. Every module of the platform is exposed as a micro-service for the end-user to experiment with. Thus for a researcher who wants to explore prompt engineering with RAG, we give them the ability to create and access the vector database, and a subsequent access to modify system prompt using programmatic access.

For example, here is a prompt created by the researcher of Antennas who wanted deeper inferences, and was ready to pass custom documents that were not in the vector database: `You are a teaching assistant. You are having a conversation with a student and the student will ask you a question. To answer the student's question use information only from the reference text that is between <Reference></Reference> and from the history of the conversation. When you answer the question, quote the text that you used to base your answer off. If you can't answer it, then say "I can't answer this question". You will add the URL for the source if it is available. You always answer the question with markdown formatting. You will be penalized if you do not answer with markdown when it would be possible. The markdown formatting you support: headings, bold, italic, links, tables, lists, code blocks, and blockquotes. You do not support images and never include images. You will be penalized if you render images. You will not wrap the output with triple backticks. Reference text:<Reference><Reference>`

[26] https://www.llama.com/docs/model-cards-and-prompt-formats/llama-guard-3/

6.11 On-boarding And Enrolling

The key challenge in provisioning LLM based access for users in any systematic structured teaching environment (e.g.students taking courses, workshops etc at a University) is on-boarding or enrolling participants and often de-provisioning them when the event concludes. All of this is currently done manually in the other available solutions like NSF funded AnvilGPT. Each user has to create their individual account and the approval and actual on-boarding will take at best 24-48 hours. Ensuring this is done systematically for a class of say 50 students will is a huge burden and hurdle even if the resource (LLM access) is provided at no cost.

AI-VERDE on the other hand , through its integration with CILogon, automatically ensures compliance by restricting access to authorized users while removing logistical burdens, allowing educators and researchers to focus on academic content without managing LLM provisioning. This vision supports equitable and efficient LLM access for the entire university community. Further, unlike other competing solutions like Arizona State University's partnership with OpenAI or Purdue's Anvil, AI-VERDE offers several additional features, including automated user on-boarding and API key management. While on-premises hosting in AI-VERDE mitigates privacy risks, it also provides proxy access to commercial models like OpenAI's GPT. In contrast, Anvil lacks extensive access management, whereas AI-VERDE enhances this by offering tools tailored for instructors to manage user access and allocate budgets to groups and classes.

6.12 Other Advantages of AI-VERDE

As mentioned earlier, while there are the 3 major categories of clients ofAI-VERDE , there are several other use cases, like enabling access to OpenAI-compatible tools (e.g., ChatboxAI[27]) , that highlight AI-VERDE 's versatile design, and ecosystem compatibility.

Maintaining budgets per course/team (e.g., billing of the usage of cloud infrastructure like AWS) and per person is important for large organizations (like universities and corporates) to ensure financial compliance. That is another advantage of AI-VERDE , i.e., its ability to track per-user usage. Our gateway model provides that abstraction using which we can track individual usage of members of a given organization. Note that currently the academic computing departments at Universities are not equipped to handle these new requirements as yet [28] and AI-VERDE is the first tool to provide such a service inbuilt.

Further, to address gaps in institutional policies, AI-VERDE team collaborated with the School of Information Department at University of Arizona, to develop a policy framework for API access, offering a scalable model for other institutions. We are hereby publishing this document at a public url [29] so that other institutions contemplating on similar ventures of incorporating LLMs (through AI-VERDE or otherwise) can use this to model their own policy documents.

7 Pilot Deployment

Since May 2024, AI-VERDE is deployed as part of a pilot project at The University of Arizona. The purpose of this pilot deployment is to stress test the system as well as understand the feasibility of its maintenance while understanding the needs of the academic community, especially in a live implementation context.

7.1 Qualitative Analysis

The current users of AI-VERDE at the University of Arizona can be divided into three major user categories.

[27]https://chatboxai.app/en

[28]we contacted Universities that have licensed OpenAI and they agree that these management functions as absent.

[29]https://tinyurl.com/llmpolicy

First, in the education frontier, AI-VERDE serves as a control plane for interacting with both commercial and open-source LLMs. Initially adopted by Faculty in AI-related courses like `INFO 555: Applied Natural Language Processing` and `BME/SIE 477/577`, it provides features like automated API key generation and budget management, with plans for expansion into non-AI disciplines.

The second major group of AI-VERDE users are researchers who analyze large collections of papers using its LLM-powered capabilities to extract insights and stay updated with fast-paced publications. For example, it was used to index 3,000 papers on antenna methodologies for the Electronics and Communications Engineering Department, enabling efficient retrieval and deeper analysis.

Third category of users were from the support departments of the University. For example, the High-Performance Computing (HPC) department utilizes AI-VERDE 's RAG-based chatbot to streamline interactions with user documentation, enhancing efficiency and improving the quality of FAQ pages.

While these are the 3 major categories of clients of AI-VERDE , There are several others as shown in Appendix A

7.2 A Quantitative Analysis

During the course of the afore-mentioned pilot deployment of AI-VERDE, 78 different users utilize it for 5 different courses and 10 research projects. The users represent a population of students, instructors, researchers and staff across multiple academic units of the university, Table 3 shows the token and API consumption statistics of AI-VERDE during a period of six months. In brief, 97 thousand API calls representing more than 110 tokens were passed through the variety of models, both self-hosted and relayed to third-party services. While these metrics are relatively low given the potential user base in an institution of this scale, especially since we are still in the pilot mode, the amount of tokens generated represents significant use and its adoption by an increasing user base and hence its huge potential for adoption in academic settings.

8 Discussion

Here we discuss the advantages of AI-VERDE specifically in a University environment.

Educational institutions should support the amazing ecosystem of commercial and on premise capa-

	Token Consumption		
	Self-hosted	*Proxy*	*Total*
Prompt	74.96M	.689M	75.65M
Completion	34.80M	.230M	35.03M
Total	109.76M	.919M	110.68M
API Calls	96,403	1,255	97,658

Table 3: AI-VERDE usage metrics during the pilot period of 05/30/2024 - 11/26/2024. In the top part of the table, the number of tokens transmitted are shown in millions. Specifically, in rows, prompt tokens represent the input data sent to various models and completion tokens represent the response data generated by the models based on the prompts. In columns, the self-hosted and proxy columns represent tokens fed and generated to models hosted in-premises, and relayed to third-party services such as OpenAI and Anvil, respectively. In the bottom part of the table, API calls represent the total number of requests made to the models either via the conversational interface and through API access.

bilities to stay in compliance. Hence the core vision and mission of the AI-VERDE team, and in turn of the Data Science Institute [30] at the University of Arizona, which created AI-VERDE, is to enable the initial hand holding and pointers required towards making a Faculty member on campus of a University, successful in using the state of the art AI for their researching and teaching resources. Specifically, AI-VERDE intends to give the equivalent of digital birth right for every member of an educational institution to have LLM access to learn, experiment and build. Further, as and when the users need higher hardware and software capabilities as part of their course or research project they will get that through the course or lab component provided in AI-VERDE .

As shown above, another core tenet at the heart of AI-VERDE , is to integrate access to the latest AI technologies while takig away the burden of orchestrating and managing the provisioning of open source LLMs. Further, from the experience of building AI-VERDE we learned that AI/LLM offering, (not just to a University setting, but to any closed user setting) is not a one-size-fits-all-solution and hence we need to mandatorily provide support to such varying levels of user skills. Hence we provide dedicated education team to provide training workshops starting from programming workshops all the way through advanced AI. Hence AI-VERDE is not just a software providing platform, but a holistic solution, where human consultations are the very first step we offer. This helps the team understand where exactly the user stands, in a compassionate and judgment-free manner.

Further, in the research frontier, Primary Investigators (PI) can bring their entitlements from other sources (e.g. NSF ACCESS, NAIR grants) and manage the access and budgets. With direct institutional log in integration via CILogon we ensure that resources access is limited to the individual and meets institutional compliance. Further, by removing the provisioning burden AI-VERDE provides educators and researchers ability focus on their core work and not have to be burdened with managing LLM access

Overall, AI-VERDE strives to provide integrate the underpinning technologies required to develop AI applications using open source software and models, and provide a seamless platform that caters to educators and researchers within an academic environment such that they can focus on developing instructional experiences and research projects without the burden of provisioning, deployment and management. Thus AI-VERDE is driving pedagogical innovation by enabling instructors to redefine their teaching methodologies. One such example included the instructor challenging students to improve upon AI-VERDE - generated answers to given homework questions, fostering critical thinking and deeper engagement with course content.

9 Future Work

In its current initial phase, AI-VERDE 's chatbot instances have been limited in scope to focus on user testing and feedback, which will guide iterative improvements. However, in the future, we plan to integrate AI-VERDE into learning management systems like D2L[31], enabling course-specific instances where Faculty can upload materials and students receive tailored interactions throughout the semester. We also aim to develop AI applications for researchers, such as a user interface for proofreading grant proposals against funding agency requirements and providing feedback to improve submissions.

Apart from the major concerns mentioned above, some minor concerns were raised in the survey. These included that chatbots currently lack the advanced reasoning ability to evaluate subjective

[30] datascience.arizona.edu

[31] https://www.d2l.com

assignments or provide meaningful feedback tied to learning outcomes. Further, it was mentioned that the chatbots are minimizing personal connection i.e., the automated interactions from chat bots undermine the forming of supportive instructor-student relationships and nuanced communications. Poor adaptability was another reason. For example, rigid chatbot capabilities cannot readily adapt coaching, guidance, and support to individual student needs and challenges. Any misunderstandings or errors from the chatbot on student inquiries or input data undermine its credibility as a knowledge source. Note that in its present form AI-VERDE does not address these challenges but is definitely part of the planned future work.

We plan to convert AI-VERDE as a chatbot assisting students with course material. This will reduces the instructor's workload, particularly on platforms like Piazza, especially in replying to rhetoric questions. Faculty can gain valuable insights into the types of questions students are asking, helping assess the class's overall understanding of the material and visibility into course effectiveness. Further, thus AI-VERDE will also provide indirect feedback on teaching methodologies by tracking student interactions. This data helps Faculty evaluate whether their teaching approach is achieving the desired learning outcomes.

Another potetial clientele for AI-VERDE are the academic and research computing infrastructure providers, like the information technology support groups. AI-VERDE can also enhance the operations of information technology based support departments. For example, the University's Information Technology Services can use AI-VERDE to build a classification and redirection system for support tickets. AI-VERDE can filter support requests, automatically categorizing them or suggesting solutions based on existing FAQs. This reduces the workload of support staff by automating the triage process, ensuring that only tickets requiring human attention are escalated.

10 Conclusion

This work introduces AI-VERDE, a platform providing LLM services tailored for academic environments, prioritizing privacy, accessibility, and adaptability. Developed with open-source components, it offers capabilities similar to commercial LLMs, with features designed for students, Faculty, and researchers. By processing all data within institutional infrastructure, AI-VERDE ensures privacy and mitigates external data exposure risks. Its integration with university systems supports research and teaching, fostering an AI-driven ecosystem that promotes innovation, collaboration, and critical thinking in higher education.

Note that even though this article provides the advantages of AI-VERDE from the perspective of its usage in an educational institution, this can be easily extended and integrated into any such similar setup, for example, employees in a corporate environment.

Ethics and Limitations

While AI-VERDE is a platform with original source code, it relies exclusively on open source models and software to host and serve LLMs. We stand on the shoulders of the research groups who graciously make their state-of-the-art models available to the general public and of the organizations who release their software freely for non-commercial use. Given that AI-VERDE exposes existing LLMs, any inherent biases in those models will be exhibited too by Verde.

One of the principal tenets of our platform is privacy-preservation. While we deliberately don't persist with any prompts or responses, we can't extend such a guarantee to any group or person who uses AI-VERDE as a proxy to manage a budget of an external platform, such as OpenAI. Also while we designed AI-VERDE to be a truly egalitarian platform, we do accept that we are still limited by the amount of dedicated hardware available in our deployment, especially from those provided by Cyverse and NSF Jetstream. However, as usage and adoption of AI-VERDE increases, possibly outside the University of Arizona itself, we expect applying for new funding to expand our dedicated hardware resources.

Acknowledgments

The authors would like to thank Nick Eddy, John Xu, Illyoung Choi, Michele Yung, Mariah Wall, Hagan Franks, Ian Mcewan, Rudy Salcido, Sarah Roberts, Paul Sarando, Maliaca Oxnam, Ajay Perumbeti, Andy Edmonds, Carlos Lizarraga-Celaya, Jeff Gillan, Megh Krishnaswamy, Michele Cosi, Jim Davis, John W, Sean Davey, Tina Johnson, Tina Lee, Tyson Swetnam and Zi K Deng without whose help and support this product would not have happened.

References

Marah Abdin, Jyoti Aneja, et al. 2024. Phi-3 technical report: A highly capable language model locally on your phone. *arXiv preprint arXiv:2404.14219*.

Accountability Act. 1996. Health insurance portability and accountability act of 1996. *Public law*, 104:191.

Joe Beda, Brendan Burns, Craig McLuckie, et al. 2014. Kubernetes: An open-source container orchestration platform. Developed by Google and maintained by the Cloud Native Computing Foundation (CNCF).

Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. 2024. The faiss library.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

David Y Hancock, Jeremy Fischer, John Michael Lowe, Winona Snapp-Childs, Marlon Pierce, Suresh Marru, J Eric Coulter, Matthew Vaughn, Brian Beck, Nirav Merchant, et al. 2021. Jetstream2: Accelerating cloud computing via jetstream. In *Practice and Experience in Advanced Research Computing*, pages 1–8.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

Evin Jaff, Yuhao Wu, Ning Zhang, and Umar Iqbal. 2024. Data exposure from llm apps: An in-depth investigation of openai's gpts. *arXiv preprint arXiv:2408.13247*.

Albert Q Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, et al. 2023a. Mistral 7b. *arXiv preprint arXiv:2310.06825*.

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023b. Mistral 7b. *Preprint*, arXiv:2310.06825.

Woosuk Kwon, Zhuohan Li, Siyuan Zhuang, Ying Sheng, Lianmin Zheng, Cody Hao Yu, Joseph E. Gonzalez, Hao Zhang, and Ion Stoica. 2023. Efficient memory management for large language model serving with pagedattention. In *Proceedings of the ACM SIGOPS 29th Symposium on Operating Systems Principles*.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. In *Advances in Neural Information Processing Systems*, volume 33, pages 9459–9474. Curran Associates, Inc.

Shervin Minaee, Tomas Mikolov, Narjes Nikzad, Meysam Chenaghlu, Richard Socher, Xavier Amatriain, and Jianfeng Gao. 2024. Large language models: A survey. *arXiv preprint arXiv:2402.06196*.

Paulo Cristiano de Oliveira, Cristiano Jose Castro de Almeida Cunha, and Marina Keiko Nakayama. 2016. Learning management systems (lms) and e-learning management: an integrative review and research agenda. *JISTEM-Journal of Information Systems and Technology Management*, 13(2):157–180.

OpenAI. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

TB OpenAI. 2022. Chatgpt: Optimizing language models for dialogue. openai.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Family Educational Rights and Privacy Act. 2014. Family educational rights and privacy act (ferpa).

Julien Simon. 2021. Large language models: A new moore's law?

X. Carol Song, Preston Smith, Rajesh Kalyanam, Xiao Zhu, Eric Adams, Kevin Colby, Patrick Finnegan, Erik Gough, Elizabett Hillery, Rick Irvine, Amiya Maji, and Jason St. John. 2022. Anvil - system architecture and experiences from deployment and early user operations. In *Practice and Experience in Advanced Research Computing 2022: Revolutionary: Computing, Connections, You*, PEARC '22, New York, NY, USA. Association for Computing Machinery.

Tyson L Swetnam, Parker B Antin, Ryan Bartelme, Alexander Bucksch, David Camhy, Greg Chism, Illyoung Choi, Amanda M Cooksey, Michele Cosi, Cindy Cowen, et al. 2024. Cyverse: Cyberinfrastructure for open science. *PLoS Computational Biology*, 20(2):e1011270.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. 2023. Llama 2: Open foundation and fine-tuned chat models. *Preprint*, arXiv:2307.09288.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Zhenhailong Wang, Shaoguang Mao, Wenshan Wu, Tao Ge, Furu Wei, and Heng Ji. 2023a. Unleashing the emergent cognitive synergy in large language models: A task-solving agent through multi-persona self-collaboration. *arXiv preprint arXiv:2307.05300*.

Zige Wang, Wanjun Zhong, Yufei Wang, Qi Zhu, Fei Mi, Baojun Wang, Lifeng Shang, Xin Jiang, and Qun Liu. 2023b. Data management for large language models: A survey. *CoRR*.

Yifan Yao, Jinhao Duan, Kaidi Xu, Yuanfang Cai, Zhibo Sun, and Yue Zhang. 2024. A survey on large language model (llm) security and privacy: The good, the bad, and the ugly. *High-Confidence Computing*, page 100211.

Defne Yigci, Merve Eryilmaz, Ail K Yetisen, Savas Tasoglu, and Aydogan Ozcan. 2024. Large language model-based chatbots in higher education. *Advanced Intelligent Systems*, page 2400429.

Yongchao Zhou, Andrei Ioan Muresanu, Ziwen Han, Keiran Paster, Silviu Pitis, Harris Chan, and Jimmy Ba. 2022. Large language models are human-level prompt engineers. *arXiv preprint arXiv:2211.01910*.

Appendix

A Full client list

These are the remaining list of courses and labs currently being supported by AI-VERDE at the University of Arizona. Specifically AI-VERDE has chat interfaces trained to reply questions related to :

1. Course content for RNR355: Introduction to Wildland Fire.
2. Research publications from the University of Arizona's Cooperative Extension [32].
3. Content from website and documentation for CyVerse: A computational framework designed to handle large datasets and complex analyses.
4. Content from website and documentation for Tech Launch Arizona [33]: Facilitating the commercialization of University of Arizona inventions.
5. Content from website and documentation MKDocs [34]: A static site generator for documentation projects.
6. Content and publications from Harwood Lab. [35]
7. Content and publications from Bonito Lab. [36]
8. Content and publications from Eller Partnership Office: [37]
9. Content and publications related to Antenna research for Hao Xin lab: [38]
10. Provides access to AnvilGPT models [39]
11. Course content for INFO 523 2024 Fall : Data Mining and Discovery.

[32] https://extension.arizona.edu/
[33] https://techlaunch.arizona.edu/
[34] https://www.mkdocs.org/
[35] https://comm.arizona.edu/person/jake-harwood
[36] https://comm.arizona.edu/person/joseph-bonito
[37] https://eller.arizona.edu/engage/partnerships-office
[38] https://ece.engineering.arizona.edu/Faculty-staff/Faculty/hao-xin
[39] https://anvilgpt.rcac.purdue.edu/

B Coda/Some Closing comments

With AI Verde, we hope to open another front of innovation in pedagogy. We hope Faculty in Universities will consider this as an opportunity to adapt to the 'novus mundus' of AI assisted learning.

In summary, what we are trying to achieve through AI-VERDE is to build the Ship of Theseus for AI technologies focusing on its use for our community. To quote what an AI researcher recently said, "AI is not a done deal. We are building the road as we walk it, and we can collectively decide what direction we want to go in, together." We think those are really wise words, and we hope that we can build an AI that really is good for humans, and not necessarily for machines themselves.