Secure Interactive Courseware Creation for Distributed Training using on-premise Generative Artificial Intelligence

Deepak Haste, Jordan Thurston, Sudipto Ghoshal
Qualtech Systems, Inc.
Rocky Hill, Connecticut
deepak@teamqsi.com,
m.renda@teamqsi.com,
sudipto@teamqsi.com

Sean Rugge, Jacob Dubois
College of Distance Education and Training
Marine Corps University
Quantico, Virginia
Sean.Rugge@usmcu.edu,
Jacob.Dubois@usmcu.edu

Jason H. Wong
Naval Information Warfare Center Pacific
San Diego, California
jason.h.wong.civ@us.navy.mil

ABSTRACT

The United States Marine Corps (USMC)' College of Distance Education and Training (CDET) has identified a curriculum need for an automated course content creation pipeline that can generate interactive eLearning products such as HTML5 Packages (H5P) from legacy schoolhouse training content such as student handouts in Portable Document Format (PDF) or Word (DOCX) format, and classroom presentations in PowerPoint (PPTX) format. This need stems from the fact that schoolhouses are resource-constrained in personnel having the requisite subject-matter-expertise (SME). Instructors are busy teaching classes with limited time to convert the static course curricula into a dynamic format within a short duration. Recent advances in generative Artificial Intelligence (genAI) has the potential to fill this void by automatically performing knowledge-extraction from the USMC training content and producing interactive content with knowledge-checks. However, cloud-based genAI models such as ChatGPT require uploading of sensitive USMC content over a network connection, thus violating USMC security policies. Consequently, there is a need for a non-cloud, non-subscription genAI capability to convert legacy content to interactive content in a standard web-ready format that is compatible with learning management systems (LMSs), thus enabling distance learning.

In this paper, the research team presents a genAI powered interactive course-package creation pipeline that uses the local processing power of laptops/desktops and offline ChatGPT-like open-source Large Language Models (LLMs) for auto-generating LMS-compliant H5P interactive learning content from schoolhouse course material within a short period of time. A web-based user-interface allows instructors to upload PPTX and Word/PDF material as contextual inputs, choose tuning-presets for creation of interactive content such as Quizzes, Dialog-Cards, Fill-in-the-Blanks and Flashcards, and review/modify the LLM responses in a human-in-the-loop interface before sending it to an automated H5P generator. The process is modular and makes provisions for swapping underlying LLMs without the need to redeploy software. It is also secure since it is laptop-based and uses on-premise genAI without requiring cloud access.

This paper documents the results of scalability tests on a standard benchmark conversion from course curriculum to H5P activity, comparing performance from various open-source LLMs on end-user desktops versus high-end GPU-enabled hardware. The paper concludes with best practices and lessons learned from these content-conversion experiments.

ABOUT THE AUTHORS

Mr. Deepak Haste is a Senior Director of Engineering at Qualtech Systems, Inc. (QSI), with a focus on customer-driven enhancements, productization, and commercialization of QSI's products through several key Small Business Innovations Research (SBIR) projects with NASA and DoD. His recent productization efforts include feature extensibility through plugin frameworks, prognostic capabilities, and embedding and interfacing of diagnostic software within onboard platforms. He is currently leading efforts to add AR/VR and Maintenance Training capabilities into QSI's product suite. Mr. Haste holds an M.S. degree in Electrical Engineering from Clemson University and a Bachelor of Technology (B. Tech.) in Electrical Engineering from the Indian Institute of Technology (IIT), Bombay, India.

Mr. Jordan Thurston is a Software Engineer at Qualtech Systems, Inc. (QSI), with focus on modernization and commercialization of learning-based tools through plugin frameworks, open-source generative Artificial Intelligence (genAI) content creation, Retrieval-Augmented Generation (RAG) pipelines, and user-centric metrics for evaluation and remediation. He currently leads efforts to integrate interactive H5P (HTML5 Package) content generation from legacy instructional formats (PPTX, PDF, DOC, XML), as well as leveraging RAG and agent-based pipelines to extract test and equipment specification data for automatic Test Program Set (TPS) generation. Mr. Thurston holds a Bachelor of Science (B.S) in Computer Science (Honors) from Central Connecticut State University (CCSU) New Britain, Connecticut.

Dr. Sudipto Ghoshal is a Vice President of Engineering at QSI, with experience in system diagnostics and prognostics, as well as design-time assessment and mitigation of operational risk and improvement of system availability. His research also involves the development of a failure-space modeling methodology leveraging general purpose system design languages such as SysML and adapting it for QSI's products. He has been a committee member of the IEEE SCC20 Diagnostics and Maintenance subcommittee since 1991. Dr. Ghoshal holds an MBA from the Kelley School of Business, Bloomington, Ph.D. and M.S. degrees in Biomedical Engineering from the University of Connecticut, and a B. Tech. in Electrical Engineering from IIT, Kharagpur, India.

Mr. Sean Rugge is a Senior Educational Technologist at Marine Corps University's College of Distance Education and Training. In his role, he is heavily involved with enterprise level LMS administration and development of the Marine Corps eLearning Ecosystem. He is an advocate for curriculum and content modernization to support training and education efforts for the Marine Corps. He holds a Bachelor of Arts from the University of Central Florida in Radio and Television Communications with a Minor in Applied Computer Sciences.

Capt Jacob Dubois is the Emerging Technologies Officer at the CDET, focusing on modernizing Marine Corps distance training and education through AI-driven solutions, content migration, and learning management systems. Capt Dubois holds a Bachelor of Science in Maritime Engineering from Massachusetts Maritime Academy and a Master's of Science in Information Technology Management from the Naval Postgraduate School, Monterey, California.

Dr. Jason H. Wong is a Cognitive Scientist with the Naval Information Warfare Center Pacific, where he supports the Office of Naval Research in improving Warfighter decision-making and human-technology teams. His government career has included measuring cognitive performance for training and simulation at the Naval Undersea Warfare Center and fostering international scientific collaboration at the Office of Naval Research Global in Tokyo, Japan. Dr. Wong holds a Ph.D. in Human Factors and Applied Cognition from George Mason University and a B.S. in Psychology from the University of Illinois at Urbana-Champaign.

Secure Interactive Courseware Creation for Distributed Training using on-premise Generative Artificial Intelligence

Deepak Haste, Jordan Thurston,
Sudipto Ghoshal
Qualtech Systems, Inc.
Rocky Hill, Connecticut
deepak@teamqsi.com,
m.renda@teamqsi.com,
sudipto@teamqsi.com

Sean Rugge, Jacob Dubois
College of Distance Education
and Training
Marine Corps University
Quantico, Virginia
Sean.Rugge@usmcu.edu,
Jacob.Dubois@usmcu.edu

Jason H. Wong
Naval Information Warfare
Center Pacific
San Diego, California
jason.h.wong.civ@us.navy.mil

INTRODUCTION

The United States Marine Corps (USMC) schoolhouses face key training and education (T&E) challenges in preparing our Marines for the future operating environment. As highlighted in the USMC "Training and Education 2030" (T&E 2030, 2023) report, there is a need to shift the current training paradigm from an industrial-age model to a more information-age learning approach via better technology integration in classrooms and courseware. To address this need, the research team devised a secure and rapid content conversion platform that produces interactive content by leveraging the power of generative artificial intelligence (AI), while still enabling the human in the loop to quality-control the content, ultimately reducing the burden on the curricula developers.

BACKGROUND

Current State of Courseware Creation

The Marine maintainers often lack access to modern training capabilities that provide up-to-date knowledge and expertise to successfully maintain systems. The current training approach in the USMC maintenance schoolhouses focuses primarily on standardized programs of instruction (POIs) including classroom components comprised of lectures with static PowerPoint slides, student handout documents, written exams, and minimal experiential learning. There is a desire for the classroom components to be richer with multimedia and interactive components. However, creating these from scratch or converting static text or images into interactive training is time-consuming. This absence of personalized and interactive training results in maintainers lacking critical, up-to-date skills, leading to decreased equipment readiness and mission delays. Moreover, this traditional classroom training model puts a burden on schoolhouses with increased training costs and inefficiencies due to outdated tools and manual processes. Specifically, the College of Distance Education and Training (CDET) has identified the following gap in content generation: *"There are a lot of PDFs, PowerPoint presentations, or Word files present in the schoolhouse curricula. While there is a need to produce interactive content - such as H5P (HTML5 Package) content – from static content, very few schools have the personnel with the knowledge and time to update the content. Curriculum content authoring capability that helps cut down the time it takes to convert or supplement the extensive amount of existing materials is extremely helpful."*

Present Needs of the USMC Schoolhouses

In order to address the aforementioned drawbacks of the present-day training practices, the USMC has called for a modern content generation pipeline that efficiently generates interactive training content for Learning Management Systems (LMS, 2025) such as Moodle (Moodle, 2025) which is being used in USMC schoolhouses. Modern LLMs leveraging generative AI have the potential to automate some of the content creation process by accepting inputs from a variety of training sources and formats and automatically outputting interactive training in eLearning format suitable for LMSs. This would lead to efficiency gains in content creation and course generation.

Specific areas where course curricula creators/designers are looking for LLM assistance are: (1) interactive content generation assistance for creation of multimedia and/or interactive learning aids that include animations and videos; knowledge checks such as quizzes and flashcards to integrate with LMS; (2) conversion to bring "legacy" content such as student handouts, presentation slides, images, movie infographics and complex diagrams into the modern

eLearning environment; (3) improve student performance through adaptive learning; and (4) advise personnel on how to best build instructional exercises in their eLearning environment to aid in enhancing course content.

Considerations for using Generative AI for USMC Content Generation

While LLMs have the potential to solve some of the aforementioned challenges, the use of commercial grade generative AI for the USMC schoolhouse use-case comes with its own set of challenges:

- USMC schoolhouses often train on sensitive systems such as combat radios for the Marine Corps Communication Electronics School (MCCES). Uploading the associated training documents to a cloud AI for processing would violate USMC IT security policies. Hence, in order to address cybersecurity requirements, an on-premise installation and deployment of the generative AI is needed to get around the need for cloud access.
- Cloud-based subscriptions from commercial vendors have prohibitive licensing and usage fees, especially if they are used for extended periods of time. Hence, the generative AI solution needs to support open-source LLMs.
- Employing generative AI to assist in the creation of interactive course content comes with its own risks of placing trust on the veracity of the LLM responses. Therefore, it is critical that there be some human-in-the-loop verification in conjunction with the automated generation of content to mitigate the effects of LLM hallucinations.
- The field of generative AI is beset with obsolescence due to the rapidly evolving nature of LLMs. An LLM could be become outdated within a matter of a few weeks or months. Thus, the content generation pipeline needs to be able to adapt to frequent updates in LLMs and be able to plug into new models.

On-prem Large Language Model (LLM)-aided Instructional Content Generation

Taking into considerations the USMC needs and the constraints on the content generation pipeline, the research team developed an AI-enabled LLM application for use by the USMC Training and Education enterprise that semi-automatically creates interactive content from course material such as Word, PowerPoint and PDFs in an efficient manner. The LLM App is non-cloud and non-subscription based, meaning that it can be hosted on an end-user laptop without needing high-end GPU-enabled hardware, with reasonable response times. The generated content can be hosted on modern e-learning systems. The App will enable a human-in-the-loop to verify, modify, and add to the AI-generated content. Coupled with learning science, the App enables tracking and analyzing of student performance in knowledge checks using Learning Record Stores (LRS, 2025). Any subpar performance can be tracked and appropriate remediation measures can be put in place to close the gap. The value to the USMC curricula developers is that it would improve their productivity by providing interactive and adaptive training tailored to individual student's skill levels, and reduces training costs by automating the creation of content that can readily integrate with LMSs.

Improvement Over the Current State of the Art

Recent advances in the integration of genAI with LMSs (Moodle AI Subsystem, 2025) only go so far as to integrate a prompt interface inside an LMS course for conversational sessions with AI, something that the course designer could do outside of the LMS. While this capability saves the effort of manually transferring the genAI responses into the LMS, the curricula developers still have to spend time and effort to transcribe the responses and manually create the interactive learning packages in the LMS. The research team's LLM App goes a step further by supporting templatized genAI prompts saved as reusable "presets" and automatically translates the genAI responses into finished H5P and SCORM products. The key innovative aspects and tools of the solution will benefit the USMC in the following ways:

a) **Reduced course creation time:** The application enables automated conversion of legacy course content to interactive content with pre-created course structure in LMSs, resulting in a potential reduction of up to 50% in the course creation time. There is no need for manual authoring of content in the LMS's user interface.
b) **Deterministic Outputs:** Since the LLM App will always have a human-in-the-loop to verify and rectify genAI outputs, the resulting products will be less impacted by LLM hallucinations.
c) **Adaptive to any LMS:** Since the LLM App will produce interactive course content adhering to industry-standard eLearning formats such as SCORM, H5P, etc., it will adapt to any LMS.
d) **Built-in Learner Analytics:** The auto-generated H5P interactive content has embedded eXperience API (xAPI, 2025) learner records that enable instructors to access student analytics and discover learner deficiencies in LRSs.
e) **Future proof:** The LLM App supports system extensibility and modularity through plug-and-play of new and updated versions of open-source AI models without the need to redeploy software. It can also reach out to on-

prem/intranet AI (DoD Cloud) or external cloud AI.

f) **Low or zero cost:** There are no licensing or addon costs since the entire software is comprised of open-source libraries. The solution also lends to low infrastructure costs since it can use standard laptop equipment instead of expensive cloud-based AI subscriptions.

g) **Ease of interactive content creation:** It enables zero-code and intuitive interactive content creation environment.

h) **Time savings:** The LLM App can extract knowledge and bootstrap the course, and convert PDFs, PPTXs, etc. into interactive content such as H5P, thus providing a significant savings of time and effort.

i) **Cybersecurity compliant:** It is secure since it adheres to USMC guidance of not reaching out to cloud-based AI.

j) **Domain Agnostic:** The LLM App is not tied to any specific learning system and can therefore be adapted across varied training domains.

SOLUTION OVERVIEW

Concept of Operations (CONOPS)

Figure 1 shows the CONOPS of the interactive course content creation ecosystem. While the approach was mostly focused on USMC Moodle, it can be applied to any LMS. The solution flow labeled by numbers is described below:

1. A Moodle plugin developed by the research team enables automatic conversion and import of schoolhouse student handouts in Microsoft Word (.docx) format into Moodle lessons within a course structure. The team also developed a utility to convert classroom training presentations in PowerPoint (.pptx) format into a Shareable Content Object Reference Model (SCORM) package. This SCORM package can be hosted in LMSs and renders itself as a clickable slideshow.

2. In the next step, the LLM App sources the same legacy content from Step 1. to apply genAI for creating interactive content in H5P format for LMSs to promote distance learning. The LLM App adheres to an open-platform architecture where it can connect to a local instance of an LLM running on a laptop or PC, or connect remotely to a high-end GPU-enabled server in the local network or in the cloud.

3. Performance metrics and scores from student activities in the eLearning system are formatted as xAPI statements. These xAPI statements are logged into the LRS as learner records. Analytics from the LRS are mined to identify learning deficiencies in students and to tailor individualized training. Based on the performance gaps, the Moodle plugin automatically generates student learning remediations/recommendations for the instructor, and assigns appropriate learning activities, familiarity exercises, and exams for the students. Thus, learning trends from LRSs are continuously being used to improve the training.
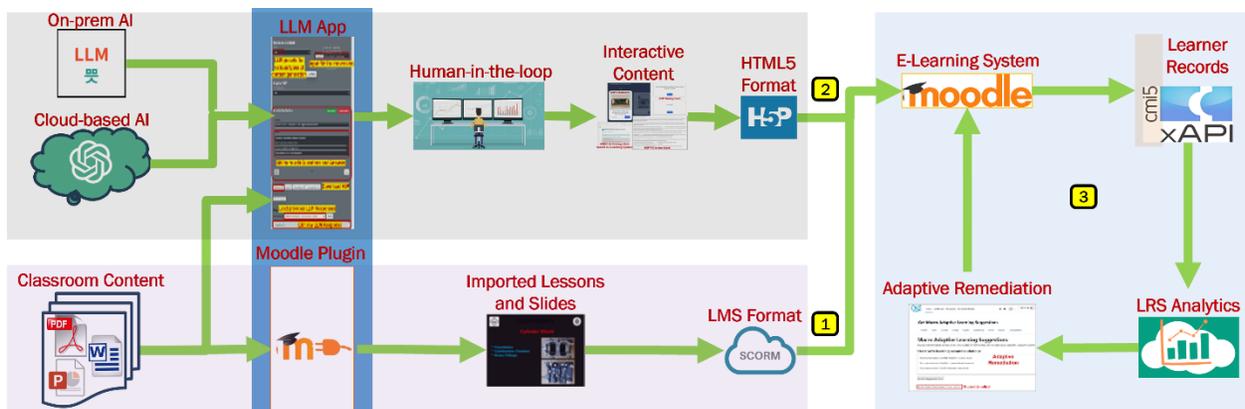
Figure 1: Solution Overview of LLM-based Course Content Generation from Training documents

A deeper dive into the individual technologies and features follows in the sections below.

Automated Lesson and Presentation Creation in LMS

The research team developed a content conversion capability to automatically import student handouts in Word (DOCX) format as Moodle Lessons (Figure 2). A Moodle plugin activates a user-interface for uploading documents

to a course. The lesson is created in a new section in the target course. The team also developed a web-interface to automatically convert classroom presentations in PowerPoint (PPTX) format to HTML pages which are wrapped inside a SCORM package. The SCORM package renders as a slideshow (Figure 3) when deployed to an LMS.
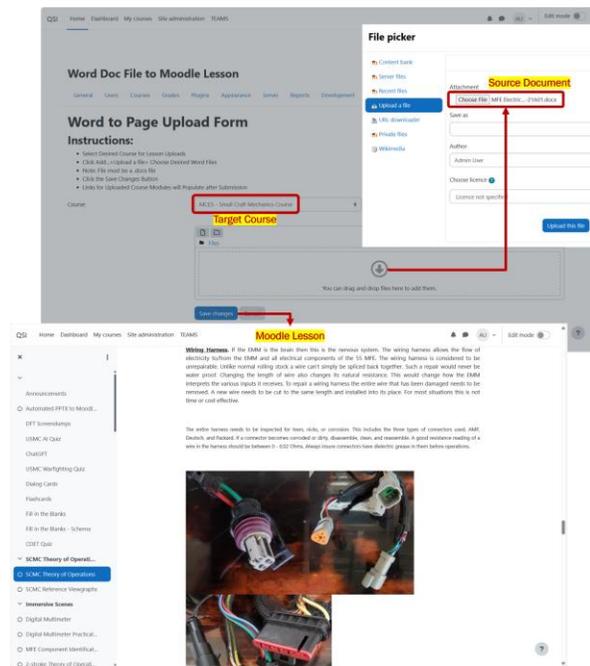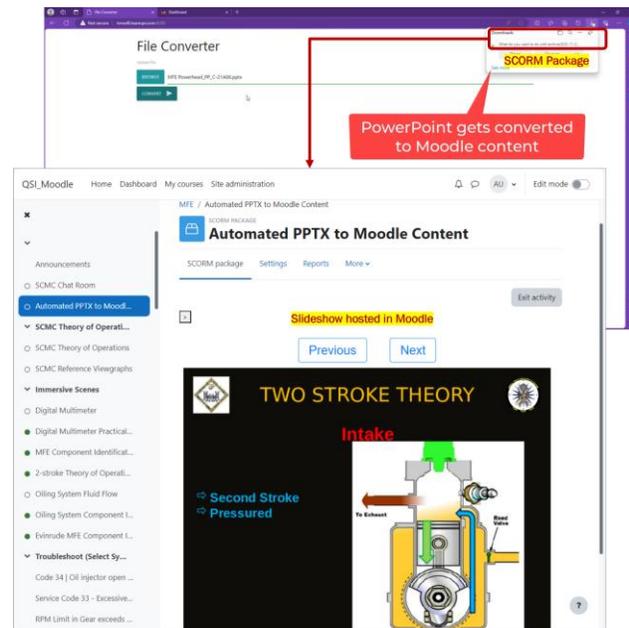

Figure 2: Word Handout to Lesson Conversion

Figure 3: PowerPoint Classroom Presentation to SCORM Slideshow Conversion hosted in Moodle

Automated Interactive Content Generation Pipeline

The content creation pipeline described in the previous section converts course content into eLearning content without modifying the underlying format or style. LLMs have the potential to take the learning content to the next level by analyzing the course material and generating rich content and knowledge checks, thus leading to a more interactive learning experience. The LLM App shown in Figure 4 enables users to upload multiple files in several formats such as PDF, PPTX, DOCX and TXT to be used as contextual input for LLM-based conversion. The user can choose the page range in
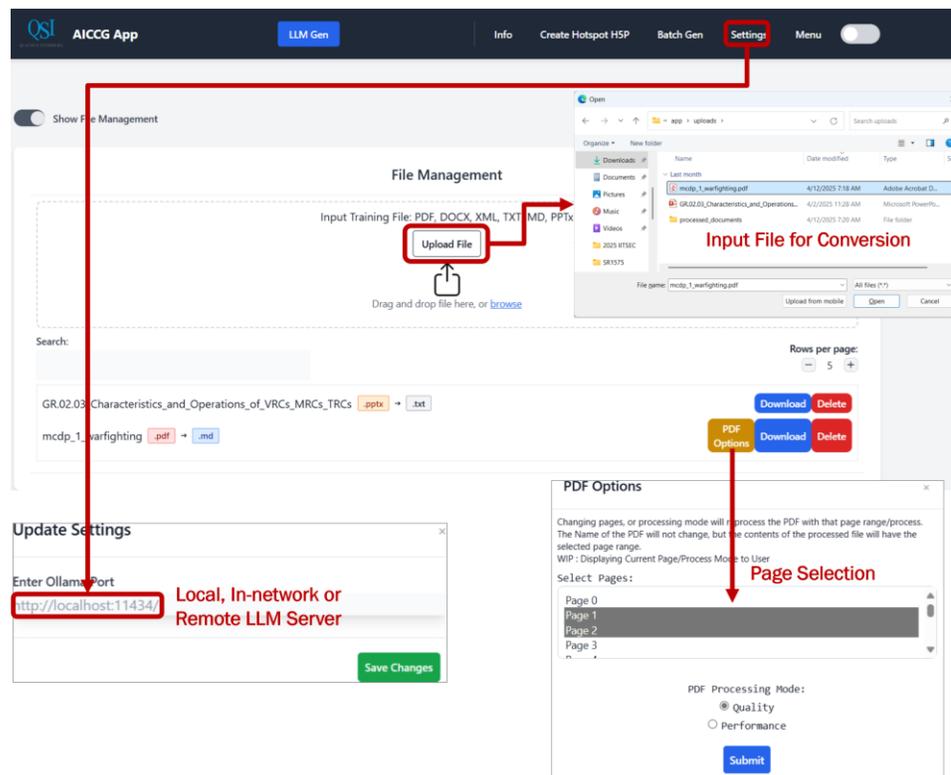

Figure 4: LLM App File Management to upload training documents for conversion

the course material to be processed by the LLM. Furthermore, the LLM can process files in two optimization modes: (a) "Quality" mode, where images, and any embedded text therein, are processed via Optical Character Recognition (OCR); or (b) "Performance" mode, where processing of text in the images is skipped. The App connects to an LLM server that processes the generative AI requests. The server connection can be configured to point to a native server running on the local machine, a server deployed within the intranet, or a server in the cloud. The App also supports batch scheduling of multiple content-conversion jobs.

The LLM App also allows the users to set prompts for creation of various H5P products (Figure 5). These parameters are stored as tuning presets for each product such as a Quiz or a Flashcard. Following is an example of a prompt stored inside a Quiz preset: "Create a quiz with 5 questions related to maintenance for electrical systems". These prompts can be customized further and saved by the user. Special-purpose vision-based LLMs such as LLaVA (LLaVA, 2025) process images contained in course files such as PowerPoint presentations, and extract descriptive texts and generate knowledge-checks from them. The user-presets also include settings for LLM servers to connect the App to. This enables the App to connect to new and more efficient LLMs as they are released, thereby making it future-proof.
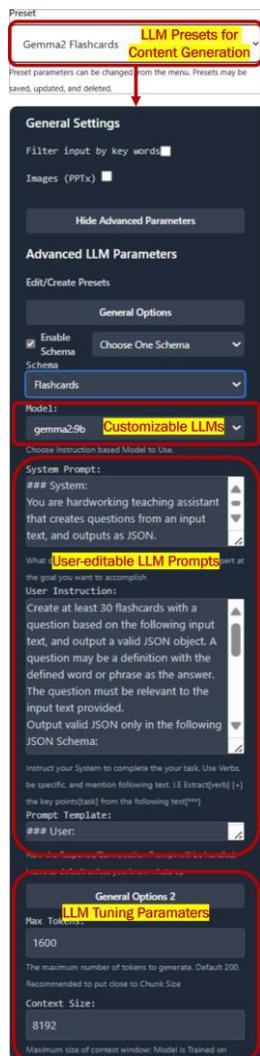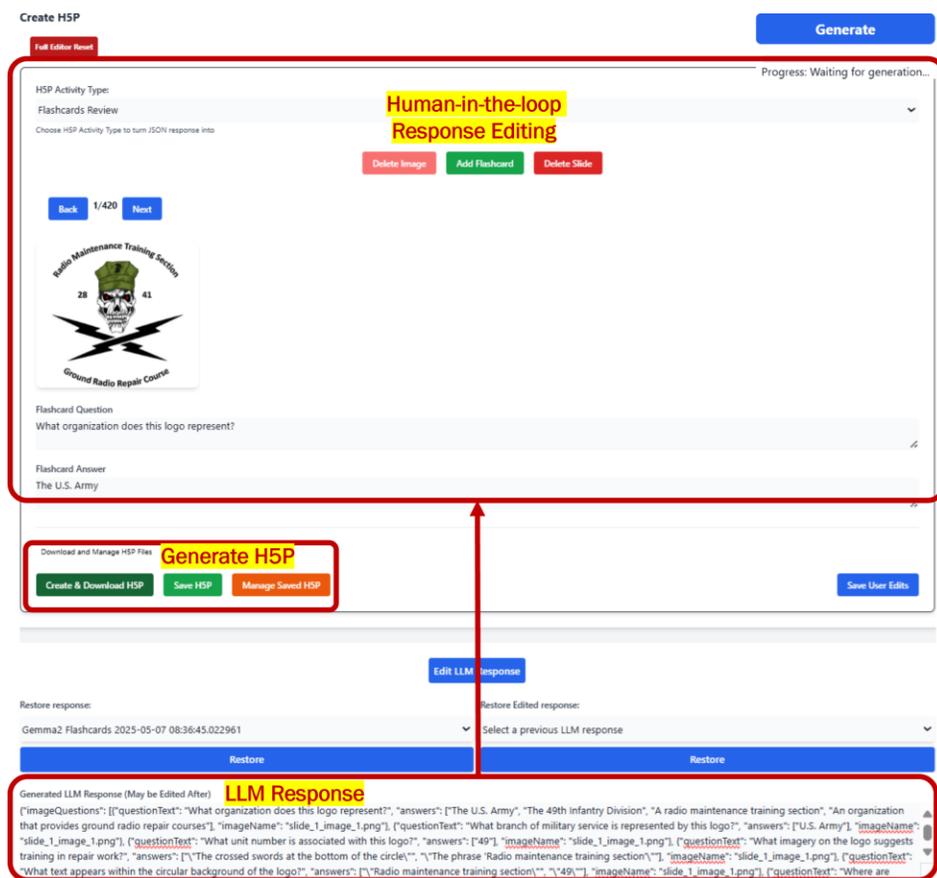

Figure 5: LLM Presets

Figure 6: Human-in-the-loop editing of LLM Responses for H5P conversion

The App displays the responses received from LLM servers in a user-friendly human-in-the-loop editor for further review (Figure 6). Furthermore, the LLM responses are converted into a human-readable format and displayed in a form. The user is able to add, edit or delete the questions and answers from the form. The user can also add, remove and/or replace images contained in knowledge-check products such as Flashcards. User edits to the LLM responses

can also be saved and retrieved later. The LLM App converts the raw LLM responses into H5P's JavaScript Object Notation (JSON) format. Once the users are satisfied with the contents of the preview, they can click on a button to generate the H5P package. These H5P packages can be readily deployed in common LMSs. The App currently supports products such as Quiz, Question Set, Fill-in-the-Blanks, Dialog Cards, Flashcards and Course Presentations.

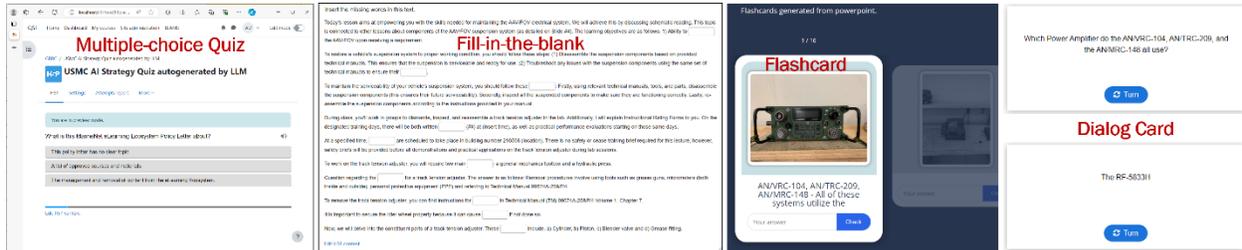Figure 7 shows some generated H5P products such as Quizzes, Fill-in-the-blanks, Flashcards and Dialog Cards.

Figure 7: Various H5P products supported by the LLM Application

Adaptive Training and Remediation

The Moodle Plugin developed by the research team integrates the built-in LRS analytics dashboards into the Moodle LMS. This enables the instructor to gauge student learning gaps in various exercises and to recommend the remediation measures, as necessary. xAPI statements embedded in the H5P activities are packaged into student learning data and stored in the LMS. Data captured includes student id, exam duration and score. The LMS, in turn, logs the student data in its data-store and periodically forwards the xAPI statements to the LRS to be used for generating analytics and to facilitate adaptive learning. Figure 8 shows the LRS analytics dashboards generated from student activities. The student metrics generated in the LRS are then used to flag performance outliers. An Instructor dashboard in the LMS identifies the students that are falling behind and recommends remediation measures (Figure 9). Upon confirmation by the Instructor, the Moodle Plugin auto-enrolls the flagged students in remediation paths and updates the LMS learning plans accordingly.

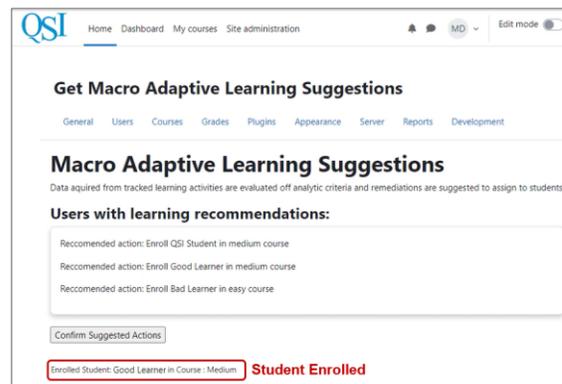**Figure 8: Student Analytics Dashboards in LRS**

Figure 9: Adaptive Remediation in LMS

EXPERIMENTAL TESTING OF LLM CONTENT GENERATION PROCESS

The research team conducted demonstrations and experimentation to assess the system extensibility and performance. The team collected usability impressions from USMC schoolhouses and CDET stakeholders. The team also assessed objective metrics of time and effort to convert courses into interactive learning content.

Interactive Content Generation Use-cases

The research team conducted interviews with CDET stakeholders and various schoolhouses to discuss and understand

the types of interactive content desired from the course material. Based on their feedback, the research team included the requested H5P products in the content generation pipeline. The team also installed the LLM App software on ONR-provided laptops and distributed them, with accompanying documentation, at various schoolhouses and at CDET. The team also provided training to familiarize the instructors and curricula developers with the software. When a facility, such as CDET, had Wi-Fi, the research team connected the LLM laptops to the network and performed remote installations/upgrades of the LLM App via a Microsoft Teams session. In other cases, the team visited the facility to install the software in-person. The USMC curricula developers used the generative AI technology to transform static educational materials in standard PDF, PPTX and Word formats into interactive H5P microlearning products such as quizzes. A summary of interactive content generated by USMC stakeholders is shown in Table 1. More details of the generated content is described in the subsequent sections.

Table 1: Summary of Interactive Content generated at USMC facilities with the LLM App

Customer	Course	Generated Content	Generation Method
College of Distance Education and Training (CDET)	USMC AI Policy document, USMC Warfighting document	H5P Quizzes auto-generated from manuals	Large Language Model (LLM) based generative AI
Marine Corps Engineer School (MCES)	Carpentry	H5P Quiz auto-generated from manual	Large Language Model (LLM) based generative AI
Marine Corps Combat Service Support Schools (MCCSSS)	Automotive Maintenance Technician Basic Course (AMTBC) Hydraulics Course	Auto-imported Moodle lesson from Student Handout	Moodle plugin to Import Word documents
	Automotive Maintenance Technician's Career Course (AMTCC)	H5P Dialog Cards auto-generated from PowerPoint classroom presentations	Large Language Model (LLM) based generative AI
	ANNEX2 Course	H5P Quiz auto-generated from Student Handout	Large Language Model (LLM) based generative AI
	AMTBC Power Plants Course	H5P Flashcards auto-generated from PowerPoint classroom presentations	Large Language Model (LLM) based generative AI
	Electrical Theory Course	Moodle slideshow from PowerPoint presentations	PowerPoint to Shareable Content Object Reference Model (SCORM) conversion web-app
		H5P Flashcards from PowerPoint classroom presentations	Large Language Model (LLM) based generative AI

Interactive Content for CDET

CDET ran the generative AI process on various USMC training policy documents, such as USMC AI Policy (USMC AI, 2024) and USMC Warfighting Manual (USMC Warfighting, 1997), to create interactive multiple-choice quizzes. Based on a feature request from CDET, the LLM App was updated to output the quiz in AIKEN format. LMSs have the means to natively import quizzes in AIKEN format, which is preferred by some curricula developers. The generated learning products will be deployed to CDET learning life cycle management course and will be accessible to trainees on their MarineNet Moodle-based eLearning platform.

Interactive Content for MCCSSS

The research team visited the MCCSSS facility at Camp Johnson, NC, and trained the MCCSSS instructors on LLM-based H5P interactive content generation. The team supplied several LLM laptops to the instructors to enable them generate their own content. The instructors converted various POI material, comprising of multiple course annexes, to Moodle courses and H5P interactive content within a matter of few hours. A sampling of the interactive content is as follows:

- Student Handout in Word format from the AMTBC Hydraulics course was converted to a Moodle Lesson by the Moodle Plugin. All formatting, including embedded images and references, is preserved in the imported Lesson.
- A PowerPoint presentation from the AMTCC was converted to H5P Dialog Cards. Dialog Cards help learners to understand various terminologies and concepts. The front of the card carries a hint, question or an image for a knowledge-check item. By flipping the card, the corresponding item description or answer is revealed to the learner.
- A student handout from the ANNEX2 Course was converted to a multiple-choice quiz. While the LLM App has the human-in-the-loop editor to edit the raw quiz content, LMSs also have the ability to edit the imported H5P quiz content.
- A PowerPoint presentation from the AMTBC Power Plants course was converted to H5P Flashcards. Flashcards have images paired with questions and answers. Learners fill in the text field in response to the questions and then check the correctness of their answers.

Interactive Content for MCES

The research team generated an H5P multiple-choice quiz from the Army Carpentry Course (Carpentry, 1995) for MCES.

Performance Testing of Open-source LLMs

The research team conducted comparative performance tests of various commercial and open-source LLMs on a common set of course material. The benchmark task was comprised of extracting text and generating a multiple-choice H5P quiz from a PDF copy of the USMC AI Strategy (USMC AI, 2024). Table 2 summarizes the benchmark results, with "Output Quality" being a subjective score based on accuracy of the quiz questions. The results show that commercial cloud-hosted LLMs such as ChatGPT, with extensive GPU infrastructure and greater number of trained parameters, fared better in terms of speed and output quality than the CPU-based LLMs running locally on laptops.

Table 2: Performance Testing of various LLMs during conversion of a PDF document to an H5P quiz

LLM for Content extraction	Nous-hermes	Falcon	Mistral	Llama	Gemma	ChatGPT
Version	2	1	0.2	3.1	2	4o
Date released	Dec-23	May-24	Sep-23	Jul-24	Jun-24	May-24
Number of parameters	10.7b	7b	7b	8b	9b	1.8 Trillion
Extraction time (hh:mm)	1:15	1:15	1:15	1:15	1:00	0:25
Output quality (out of 10)	4	4	4	4	6	8

CHALLENGES, LESSONS LEARNED AND FUTURE OPPORTUNITIES

Performance of On-Prem Hardware Setup

As shown in the previous section ("Performance Testing of Open-source LLMs"), consumer laptops have a performance shortfall due to the absence of high-end GPUs which are necessary to run the large generative AI models. Therefore, it takes considerably longer to create course content solely on the processing power of the CPUs. Moreover, low-end LLMs have a tendency to hallucinate and generate less optimal responses. Open-source LLMs are becoming more efficient in terms of their ability to perform computations on a smaller processing footprint. With each new version, a relatively small generative AI model can produce a vastly improved quality of content. Since the LLM App follows an open-architecture, it can adapt to newer LLMs in a plug-and-play fashion, and hence is future-proof.

Future Work

A significant portion of future enhancements will focus on performance improvements for course conversions that use local or intra-network LLMs. The research team will incorporate Retrieval Augmented Generation (RAG, 2025)

techniques for improved scalability and for sourcing large documents as contextual inputs to the LLMs. Additionally, the research team intends to expand the range of H5P products that can be generated by the LLM, including:

- *Videos*: (a) Use vision and speech-based generative AI to create knowledge-check questions, e.g., play a segment of narrated video clip and ask what maintenance action is being performed; (b) Generate H5P Interactive Videos.
- *PPTX*: (a) Convert slides to narrated videos (text-to-speech); (b) Create H5P Hotspots from labeled images,
- *DOCX*: (a) Convert Lessons to Branching Scenarios with knowledge checks; (b) Create Interactive Books.

The research team will plan to transition this technology to the Marine Corps Virtual Platform (MCVP). The team will also explore hosting in a secure, Government-managed cloud or hybrid Government-managed (DoD NIPRNet) environment. The tool will utilize local AI models for Distro D and cloud AI models for Distro A source materials.

CONCLUSIONS

In this paper, the research team demonstrated the concept of an LLM-based on-prem interactive content generation pipeline that provides time and cost savings by utilizing open source generative AI models and inexpensive hosting hardware. The solution provides an ability to automatically generate interactive H5P content from USMC schoolhouse training documents, thus avoiding the need to manually author the material. The resulting content tracks student performance via xAPI logs stored in LRSs, and provides analytics to assess the students' knowledge, skills, abilities (KSAs) and learning gaps, and recommends remediatory learning paths. The pipeline facilitates human-AI teaming to develop and manage training for the USMC faster and more effectively. This, in turn, can help the USMC achieve the strategic goal of modernizing schoolhouse training that reflects a cutting-edge, information-age learning enterprise.

ACKNOWLEDGEMENTS

REFERENCES

Carpentry (1995). *FM 5-426 - Carpentry*. *https://en.wikipedia.org/wiki/Learning_management_system*

LLaVA (2025). *Large Language and Vision Assistant*. *https://llava-vl.github.io/*

LMS (2025). *Learning Management System*. *https://en.wikipedia.org/wiki/Learning_management_system*

LRS (2025). *Learning Record Store*. *https://xapi.com/learning-record-store/*

Moodle (2025). *Moodle - Open-source learning platform*. *https://moodle.org/*

Moodle AI Subsystem (2025). *Unlocking the power of AI*. *https://moodle.org/mod/forum/discuss.php?d=462651*

RAG (2025). *Retrieval-augmented generation*. *https://en.wikipedia.org/wiki/Retrieval-augmented_generation*

SCORM (2025). *An introduction to SCORM*. *https://scorm.com/scorm-explained/one-minute-scorm-overview/*

T&E 2030 (2023). *Training and Education 2030*. *https://www.marines.mil/Portals/1/Docs/Training%20and%20Education%202030.pdf*

USMC AI (2024). *United States Marine Corps Artificial Intelligence Strategy*. *https://www.marines.mil/Portals/1/Publications/USMC%20AI%20STRATEGY%20(SECURED).pdf*

USMC Warfighting (1997). *United States Marine Corps Warfighting Manual*. *https://www.marines.mil/portals/1/publications/mcdp%201%20warfighting.pdf*

xAPI (2025). *eXperience API (xAPI)*. *https://xapi.com/overview*