Using GANs and LLMs to Create a Synthetic Information Wrap – Lessons of the POBL Project

D.J.H. Burden[1] and D. Keane[2]
[1]Independent Consultant, Daden Limited, Birmingham, UK. david.burden@daden.co.uk
[2] Project Manager, Vindico Limited, Cardiff, UK. dan@vindico.net

Abstract — Information Injects can provide a valuable depth to exercises but can be time-consuming to create and can struggle to reflect the events of a dynamic operation. There has been much discussion about the potential of AI, particularly Large Language Models (LLM) to address this – but how viable is it in current practice? In 2022 we reported at ITEC on our work on developing two Artificial Intelligence for Information Injects (AI4II) applications. In 2023 we secured funding to further develop this software in support of the British Army's Populating the World of Training competition, funded by Army Innovation and delivered by the Defence and Security Accelerator (DASA). The new project was called POBL. This paper reports on the work done in the POBL project in two main areas – extending the synthetic wrap to include live exercises, and the application of LLM and GAN technologies to enhance the generation of the synthetic wrap. Overall, the project suggests that LLMs can help in the generation of injects but are currently best used in a framework where they are an addition rather than a core component, and that a reversionary mode is always available.

1 Purpose

As battles are increasingly being fought within urban and urbanized environments, and as the civilians within those environments are likely to have access to social media, it is becoming increasingly important that exercises and wargames represent the information space within that environment in a dynamic way – both so that players and exercising troops can better understand their impact on the environment and civilian population, and so that they can leverage that information space for their own intelligence and effects.

Many traditional approaches to synthetic information wraps have provided a rigid or manpower-intensive approach (or both). Prior work demonstrated an innovative templating approach to synthetic information wraps for both Virtual and Constructive training, but was prior to the advent of commercially available LLMs and GANs. In order to provide a more complete, current and effective solution to the need for a synthetic information wrap, the POBL project sought to improve on the prior work in three main areas:

- Integrating Virtual and Constructive solutions;
- Extending the synthetic information wrap to Live Exercises; and
- Exploring the use of LLMs and other GANS to improve the output, and generation of, the synthetic information wrap.

2 Key Takeaways

The key takeaways from our work on using AI within POBL to date are that:

- LLMs are currently best used in support of systems rather than as the core component for inject generation;
- Image generation is best for very generic use cases, and video generation still has a way to go to be properly useful; and
- Always have a reversionary mode.

3 Introduction

In March 2023, the UK MOD's Defence and Security Accelerator launched a call for projects with the aim of "Populating the World of Training". The call had £2.845m funding, and was seeking 12-month, £100k-£300k proposals, with a TRL of 3-6. The call was part of the Audiences, Actors, Adversaries and Enemies (A3E) Human Terrain strand of the Future Collective Training System (FCTS). The call sought "*novel and innovative proposals that will help deliver a credible, complex and representative human terrain to effectively train British Army personnel and other Land Forces operating in the land environment*" and was looking for a "*a free-thinking A3E capability that delivers cues, stressors and frictions across the human, physical, environmental and information domains. This training must be seamlessly blended across Live, Virtual and Constructive (LVC) options, and be realistic, dynamic and adaptable to the developing exercise.*" [1].

Daden had previously developed two synthetic wrap systems for UK MOD, one focused on constructive wargames (ASMSW) and one focused on virtual simulation systems (AI4II), and a paper on these was presented at IT2EC in 2022. [2]

4 Approach

The new project was christened POBL – Welsh for "people" – since Vindico, our development partner is based in Wales. The approach for the project was to move the core templating system that underpinned AI4II and ASMSW onto a more modern technology platform, and enable the same platform to manage both real-time and turn-based inject generation, and then to:

- Extend the architecture to accommodate Live Exercises; and to
- Explore the use of LLMs and other GANS to improve the output, and generation of, the synthetic information wrap.

In addition, POBL was given the ability to accept and generate HLA messaging so that it could be more readily integrated with simulation federations.

Since POBL is intended to be primarily a military application all AI work was done on locally run models and software, which also helped to mitigate some of the "ethical messaging" issues of commercial LLMs.

5 Results and Discussion

5.1 The Merged System

Figure 1 shows the high-level architecture of the new POBL system.

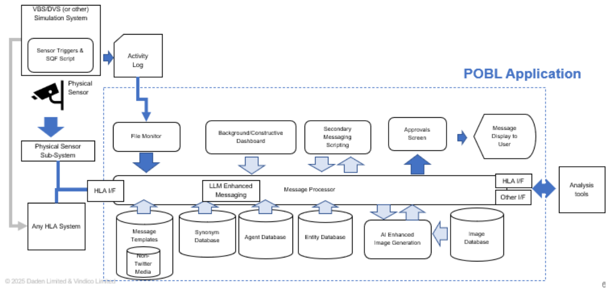

Figure 1: POBL Architecture

The templating approach has already been described [2], and a key element of the design remains the way that messages are layered by the system – but with the user only seeing the resultant complete feed. With POBL this layering includes both background messaging and those from virtual and physical triggers, allowing an exercise to be provided with any desire mixed of messages, as shown in Figure 2

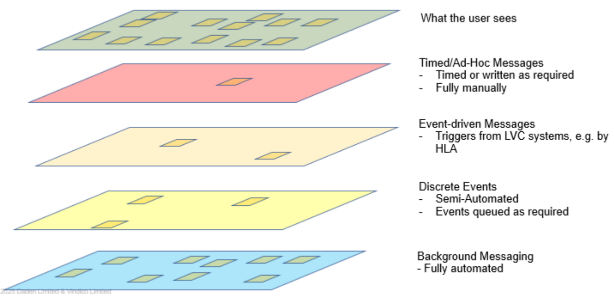

Figure 2: POBLs Layered Approach

Figure 3 shows the Exercise Dashboard screen which is at the heart of controlling the POBL system, and Figure 4 shows the Moderation screen, with generated messages – moderation being set to pre or post publish as required.

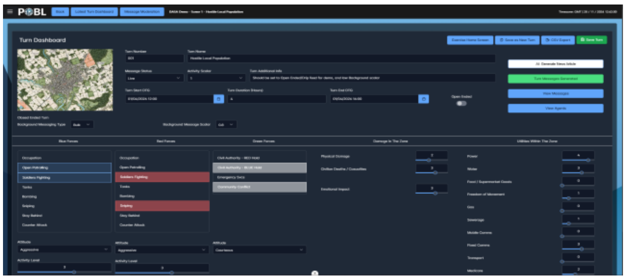

Figure 3: POBL Exercise Dashboard

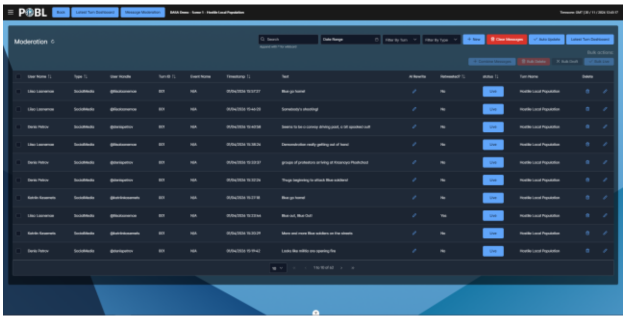

Figure 4: POBL Moderation Screen

5.2 Adding Live Exercise Support

The concept for Live Exercise support was that sensors could be placed in an urban area, and would then generate messages as through representing civilians, which would then be merged with background messages appropriate to the exercise tempo and passed back to the exercising HQ. As such a physically robust sensor system needed to be combined with image recognition software in order to generate standard entity-based triggers which could be used by POBL to generate the messages.

For the sensor a single-board micro-computer within a rugged camera trap case was used, and equipped with PIR (Passive Infrared) sensors for motion detection. Cellular or wifi (if available) could be used for the back-haul. Other technology, such as LoRaWAN, was also explored, but did not have the desired range. For image detection we implemented TensorFlow-based object recognition models to classify detected entities in real-time. All data was processed through edge computing, ensuring rapid and localised decision-making. The result was a JSON trigger message which could be fed into POBL as any other external trigger message.

Figure 5: The Physical Sensor system

Two main challenges were encountered in developing the system:

- The limited processing power of edge devices occasionally created delays in the TensorFlow-based object recognition.
- The baseline object library for object recognition constrained the system's ability to classify and react to a diverse range of real-world entities. This required development of a custom object library to enhance performance, and to enable future support in military training scenarios

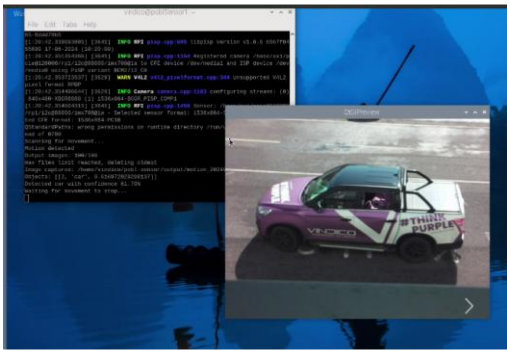

Figure 6: Entity detection

The core result, though, was that the system enabled real-time message generation by POBL based on detected physical activities, proving the concept and helping to potentially bridge the gap between Live and Virtual/Constructive environments, ensuring seamless training integration.

5.3 Using AI to Enhance Messaging

We used Llama 3, a community-driven LLM, as the main text LLM for POBL. The model's capabilities were optimised to produce outputs aligned with product and exercise objectives. The LLM capability was typically accessed through "helper" buttons on the POBL screens, and could generate:

- New templates based on a prompt;
- Additional synonyms based on existing synonyms;
- Avatar profiles – leveraging existing agent parameters (e.g. gender, age, allegiance, nationality/faction);
- Mainstream media (MSM) news reports by aggregating individual messages.

In addition, there is a dedicated LLM chat screen where the operator could try out ideas and generate extra content.

This use of LLM proved highly successful, reducing operator time on tasks (particularly profile and MSM report generation) and also helping to enrich the messages being generated. However even with a local LLM there were still some issues of the LLM's "ethical" filter preventing the generation of some messages, and some prompt engineering was required to circumvent this.

One learning was that as prompts were developed many elements could be incorporated into the "exercise pre-prompt" so as to avoid the need for the operator to be explicit each time.

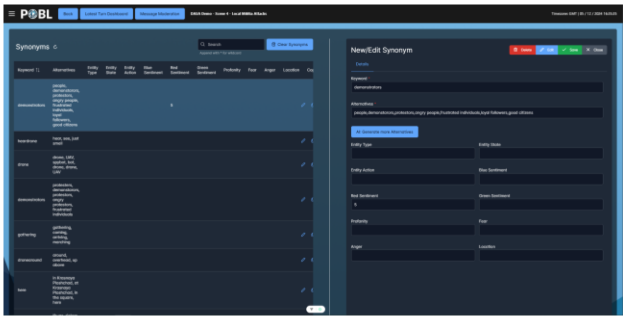

Figure 7: An AI /LLM "helper" button on the synonyms page.

5.4 Using AI for Still Imagery

The use cases for generating still imagery within POBL (or any synthetic information wrap) include avatar profile pictures/avatar, "smartphone" pictures, MSM imagery and other inject pictures to help set the context and detail events. For POBL we used Fooocus AI (sic), to create high-quality synthetic images based on user prompts or system data. Avatar profile pictures were readily generated (similar to the thispersondoesnotexist.com website), but the generation prompt could be populated with information from the agent biographic data (e.g. gender, age, race) and the scenario setting. The generation of other imagery was passable, but often required significant prompt engineering to get just right – rather than just providing the AI with a message and asking it to generate an image – particularly given the brief nature of many social media posts. In particular, it could be hard to get images which were consistent across the whole exercise, and also to get consistency with what players may be seeing through other systems (e.g. VBS4) and to match the look of some potential geographical locations and even specific equipments. Again, the use of scenario data could help to tune the pre-prompt to improve consistency.

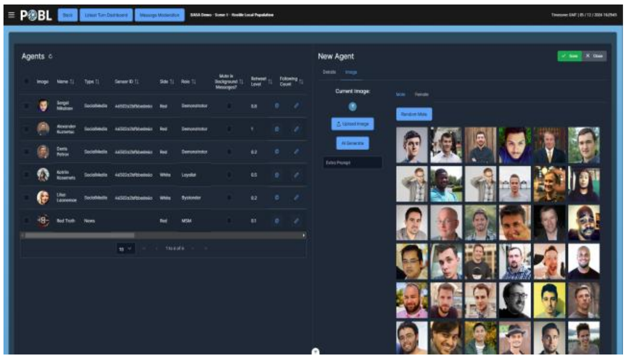

Figure 8: GAN generated avatar images

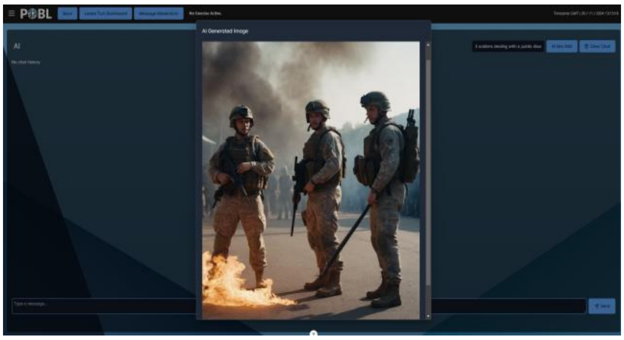

Figure 9: GAN generated image for the prompt "3 soldiers dealing with public disorder"

5.5 Using AI for Video Imagery

The use cases for video imagery are similar to those for still imagery, and include social and mainstream media video, propaganda video and inject video. After investigating a number of different frameworks and pipelines we settled on the open-source ComfyUI framework (https://www.comfy.org/en/). The workflow for this is shown in Figure 10 and Figure 11.

Figure 10: ComfyUI Workflow - Stage 1

Figure 11: ComfyUI Workflow - Stage 2

It consists of text-to-image stage followed by an image-to-video stage. However, there were similar, but more extreme, problems in video compared to still image generation around consistency and prompt engineering. There was also a significant additional overhead in generation time. As a result, the video generation did not provide adequate assets within the near-real time conditions in which POBL is expected to be used in. However, this is an area where technology is improving on a monthly basis, and improvements in this area are being monitored and incorporated into POBL as they become usable within the POBL context.

The project also looked at some of the commercial video generation offerings, more for the pre-preparation of video injects than for dynamic creation of content during an exercise. As one example the initial Populating the World of Training included a graphic novel style storyboard of a potential scenario, including an image of a newsreader from the hostile power. It was a simple task to upload this image to the Heygen site (https://www.heygen.com/) , provide a short news script generated by POBL and then have Heygen provide an animated video of the newsreader reading the script (Figure 12).

Figure 12: Heygen generated video from a still image

5.6 Using AI to Directly Generate Messages

The final objective in the project was to assess the extent to which an LLM could be used to completely bypass the template/synonym model used by POBL (and its predecessors). Within POBL the current exercise state is represented by around 20 variables (as shown in Figure 3) which include the levels of military activity/fighting for Red and Blue, the levels of damage to civilian infrastructure (power, water, sewerage, medical, transport, housing) and the level of civilian casualties. Each agent also has parameters around their levels of fear, anger, profanity and sentiment.

As shown in Figure 13, in this mode the current state parameters were used to automatically generate a prompt. This prompt could then be used to generate the messaging – either as individual social media post or as longer MSM posts. This could then be combined with the Fooocus AI or ComfyUI workflows to generate supporting still or video imagery.

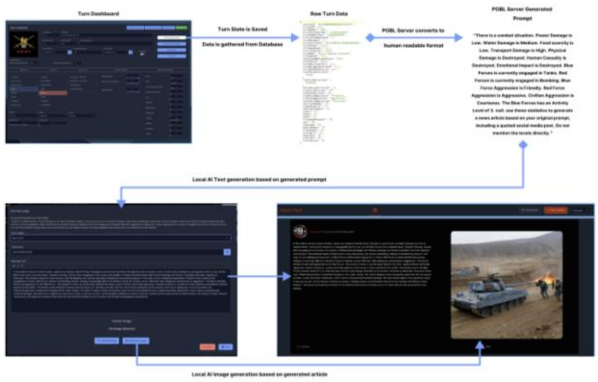

Figure 13: The LLM Direct pipeline

In testing we found that whilst this approach could provide reasonable MSM stories, it could not generate the variety of messages which were required for social media – although further refinement of the prompt engineering may improve this.

However, a more effective approach was found to be a hybrid one, where POBL generated messages through its standard template approach, but the LLM then rewrote them in a variety of different styles. This not only gave more appropriate and consistent messages, but also meant that the number of templates and synonyms that need to be created in POBL can be dramatically reduced as the LLM both provides the variety (which would otherwise be generated through synonyms and template variation) and the personalisation (by changing the wording for messages based on each agent's personality profile). One challenge though was in how the LLM incorporates context dependent synonyms, such as how it can substitute different terms for Red and Blue forces and supporters. We are further exploring this hybrid approach to make it a core part of the POBL offering.

6 Lessons Learned

In working with a variety of AI systems as part of the POBL development a number of challenges were encountered and lessons learned. These include:

- Whilst POBL leveraged Llama 3 for text generation and Fooocus AI for image outputs, these approaches needed to be assessed and fine-tuned for each scenario. The limitations in pre-existing training data required careful adaptation and prompt engineering.
- Prompts had to be iteratively refined, and we had to introduce placeholders (e.g., {@sentiment}) to dynamically adapt content based on real-time data.
- It was important to implement middleware for efficient data processing and caching, optimising response times without compromising quality.
- Workflows had to be designed that ensured consistency across media modalities, but the lack of mature, defence-ready video-generation technology limited immediate applicability.
- Ethical and security concerns were mitigated by hosting the AI locally within a closed network, ensuring secure data handling and ethical oversight of outputs
- Whilst hallucinations still occurred they were less critical in a social/media context as they could just represent natural "noise" and confusion. However, the POBL moderation screen provided for pre or post publishing validation and correction
- Models run locally with pre-configured hardware and caching techniques could manage the high computational demands, but scaling to larger scenarios may require additional resources and higher specification equipment.

7 Future Work

POBL is now at TRL 6/7. We are continuing to refine the use of AI within all areas of POBL, but particularly in video generation (where the biggest challenges remain) and in a hybrid message generation approach (where the biggest efficiency improvements are probably to be found). We are also considering both how to make the agent population more dynamic (such as in fear/anger/sentiment and followship), and even interactive – although this may be better achieved through the integration of POBL with existing "influence range" type applications.

8 Conclusions

The success of POBL in reaching TRL 6/7 shows that it is possible to develop a single system to generate a synthetic information wrap for virtual, constructive, wargaming and even physical training – potentially saving costs and time, and increasing consistency across the training estate. It has also shown that AI can be a valuable tool in helping to enhance agent profiles, in generating messaging (social and MSM) and images, although near-real-time-video is not quite there yet. Some of the other challenges of AI also suggest it is better, at the moment, viewed as a support tool rather than an essential part of the pipeline. In summary we consider that POBL is well placed to address the challenge initially presented of "populating the world of training" in a very cost-effective and flexible manner.

References

[1] UK MOD/DASA (2023). *Competition: Populating the World of Training (Phase 1)*. Gov.uk. https://www.gov.uk/government/publications/competition-populating-the-world-of-training-phase-1.

[2] Burden, D.J.H. (2022). *Automated generation of information injects.* IT²EC 2022. London, UK.

Author/Speaker Biographies

David Burden has worked in AI and VR since the 1990s, delivering over 100 projects through his company, Daden Limited. David is also currently series co-editor for Taylor & Francis on The Metaverse Series. David is an ex-Royal Signals officer and is studying for a PhD in Wargaming Urban Conflict.

Dan Keane is Chief Project Officer at Vindico ICS. With extensive experience in defence, technology and project management he delivers solutions for a diverse portfolio of clients. His work includes AI-driven solutions for military training across Live, Virtual, and Constructive environments, providing effective and adaptable tools for complex operational needs.