N252-109: Automated Intelligence Preparation of the Battlespace from Open-Source Information

ADDITIONAL INFORMATION

N/A

TECHNOLOGY AREAS:

None

MODERNIZATION PRIORITIES:

Advanced Computing and Software | Human-Machine Interfaces | Trusted AI and Autonomy

KEYWORDS:

Artificial Intelligence, Machine Learning, Mission Planning, Intelligence Preparation of the Battlespace, IPB, Marine Corps Planning Process, MCPP, AI/ML

OBJECTIVE:

Develop automated tools using artificial intelligence/machine learning (AI/ML) to generate inputs to Intelligence Preparation of the Battlespace (IPB) for the Marine Corps Mission Planning Process.

DESCRIPTION:

The Marine Corps Planning Process (MCPP) requires Marines to perform an IPB, which is the "systematic process of analyzing and visualizing the portions of the mission variables of the adversary, terrain, weather, and civil considerations in a specific area of interest and for a specific mission. By applying IPB, commanders gain the information necessary to selectively apply and maximize operational effectiveness at critical points in time and space." [Ref 1].

However, the IPB is a time and manpower intensive process requiring Marines to analyze physical, temporal, cognitive, and virtual considerations [Ref 2]. Open-source information is often valuable input to the IPB given the dynamic nature of many IPB considerations, but this information is frequently the costliest to find and incorporate. In a tactical situation, Marines may not have the time or expertise to perform a full analysis across these domains.

Current IPB inputs are also not easily actionable by AI/ML analytics due to the analog nature of the resulting products. Automated generation of the IPB in structured, computer-readable formats would enable additional AI/ML tools to incorporate and exploit those products in other parts of the MCPP. A tool that solve this problem should scan through open sources of information of IPB-relevant content. Open sources to include scans should be both user-provided and discovered by web scraping. It should then retrieve and parse that information into a structured, computer-readable format usable by other analytics supporting MCPP. IPB content generated by the tool should include references to the original source of the information. The tool should estimate the accuracy of the information based on the source, the correlations between different open sources, and the correlations to authoritative U.S. sources. Marines should also be able to influence rankings of specific generated content or source sites.

The focus in developing the tool should be in discovery of relevant information to the IPB process, computer understanding of that information, AI-assisted data and cleaning transformation to enable rapid, actionable use by other algorithms, and content evaluation (i.e., source quality, accuracy, uncertainty estimates) of the generated information. Previous work in this area used Natural Language Processing (NLP) and Semantic Search techniques but were lacking in both quality and specificity of the generated information. Generally, results from previous techniques still required significant human review and refinement. Recent innovations in Large Language Models (LLMs) and Deep Learning are expected to support significant improvements in the development of a solution.

PHASE I:

Determine the technical feasibility of a concept for the automated generation of IPB content. Develop a quality metric for discovered data sources to include accuracy/uncertainty considerations and user preferences. Prepare a Phase II plan.

PHASE II:

Develop and evaluate a prototype for the automatic generation of IPB products from open sources. The prototype should use both a pre-defined list and automatically discovered sources. The prototype should generate IPB products targeting a minimum of three IPB considerations (e.g., civil considerations, adversary order of battle, adversary tactics, etc.). Generated products should be stored in data formats optimized to the structure of the underlying information such as JSON, Geo Tag Image File Format (GeoTIFF), and Network Common Data Form (NetCDF., etc. A product quality metric should be developed as an extension of the data source quality metric. The product quality metric should granularly evaluate information found in data sources for accuracy/uncertainty via correlation with other sources (both other open sources and authoritative sources) or other means. For Phase II authoritative sources need only include trusted public sources (e.g., open access government webpages).) The prototype should support user input to refine source quality and the accuracy of specific information in the system.

Produce the following deliverables: (1) a working prototype developed according to the extended Phase II requirements; (2) product quality metric methodology; (3) a test report documenting results of prototype evaluation.

PHASE III DUAL USE APPLICATIONS:

Support the Marine Corps in transitioning the technology for Marine Corps use. Develop the software for evaluation to determine its effectiveness in either a formal Marine Corps schoolhouse or other training setting. Incorporate the product into larger AI-enabled mission planning tools such as the Higher Echelon Mission Planner. Generated IPB products will be presented directly to the Marines via the planning tool and used as inputs to other AI/ML analytics support course of action analysis.

As appropriate, focus on broadening capabilities and commercialization plans. Development of affordable, scalable, non-proprietary technologies are needed to accelerate the transition of the Marine Corps to an information age model. The commercial sector is developing some of these AI-enabled technologies, but they often do not deal with critical issues regarding non-existent, limited, or low-quality source data, do not address the diversity of data modalities employed by Naval forces, and often come with prohibitive licensing and usage fees. This technology will have broad application in the commercial sector. Examples of businesses that would be interested in this technology include companies developing generative AI, companies providing search engines, and news media. All of these companies would benefit from the ability to identify low-quality source data to improve the accuracy of the information they provide to their customers. Additionally, companies focused on Generative AI have also been pushing to incorporate non-textual data into their offerings. Research with multi-modal data sources would benefit them as well.

REFERENCES:

1. “Marine Corps Planning Process, MCWP 5-1.” Department of the Navy, 24 August 2010. https://www.marines.mil/Portals/1/MCWP%205-1.pdf
2. “Intelligence Preparation of the Battlespace, MCRP 2-10B.1.” United States Marine Corps, 1 July 2023. https://www.marines.mil/Portals/1/Publications/MCRP%202-10B.1.pdf?ver=WZgANNGDKsmcphvgtPdt-Q%3d%3d

TOPIC POINT OF CONTACT (TPOC):

TPOC-1: Peter Squire
PHONE: N/A
EMAIL: peter.n.squire.civ@us.navy.mil

TPOC-2: Elias Ioup
PHONE: N/A
EMAIL: elias.ioup@nrlssc.navy.mil