

Adaptive Neural Machine Translation Solution OCR Research and Domain-Adaptation Scaling Prototype



Lilt, Inc.
Emeryville, CA
www.lilt.com

Contact:

Phil Stiefel
 Director, DoD
 Lilt, Inc.
phil@lilt.com

Topic Number: AF192-001

SYSCOM: Office of Naval Research
 (ONR) | www.nre.navy.mil

Program Sponsor: Office of Naval
 Research (ONR)

Other Potential Programs: NAVFAC
 FE, NAVWAR, NCIS, C7F, III MEF, JRMCS SRF,
 CNFK, overseas military hospitals

Current TRL:

Lilt's baseline language translation
 technology: 9
 Additional functionalities being
 researched and built through this SBIR (as
 of July 2024): 6

Projected TRL: 8 / Q4 2024

Keywords: Translation, Large Language
 Models, AI/ML, Optical Character
 Recognition (OCR)

SBIR Pavilion



2024 Navy Gold Coast | August 20 - 21, 2024

THE CHALLENGE

The Navy Information Warfare Community and overseas Navy units/installations frequently encounter paper and scanned documents that need to be translated into or from English. To accomplish translation, the documents must be converted from picture/scanned image format to digital text.

The process of converting text in scans/images is known as Optical Character Recognition (OCR). This Phase II SBIR project addresses capability gaps in existing OCR technologies, and researches new OCR advances for Navy users. Secondary goals of this project include:

- Developing faster methods to enable AI applications with large neural images to deploy in Navy-accredited security environments.
- Experimenting with scaling language translation AI model fine-tuning across large user groups.

THE INNOVATION

Lilt developed and implemented software upgrades to increase OCR quality and reduce the human labor needed to process image/scanned documents. Lilt will also conduct technical research on methods for OCR error detection, improved OCR of structured data, and enhanced OCR capabilities for images within text documents.

Lilt is also:

- Developing custom software solutions to facilitate deployment of large neural images in DoD security-hardened environments.
- Refining a process that can optimize scaling fine-tuned AI language translation models across large enterprises.

THE NAVY BENEFIT

- 1) **Faster Translation Turnaround Times:** The technology advances pioneered by this SBIR effort will reduce the time required for Navy translators to translate documents, which will provide faster/fresher information for Navy leaders.
- 2) **Enhanced Translator Productivity:** Recruiting and training additional Navy linguists is a slow and expensive process. By providing Navy translators with software tools that leverage recent advances in AI and LLM technology, the Navy can greatly increase the efficiency and output of its existing language translation workforce. This will help the Navy cover foreign language translation capability gaps in cases where translator billets are not fully filled or during conflict/crisis.

THE FUTURE

Going forward from this project, Lilt will continue to identify centers of translation capability need across the Navy/USMC. The technology developed in this effort will be made available for broader DoD use via software licensing.