WOPR, or the Web-Scale Operational Platform for Research, is a comprehensive, extensible big data environment with integrated machine learning. It features a highly scalable object storage-based data layer, a flexible, high throughput data pipeline, and an end-user interface for data discovery, curation, and exploitation across disparate datasets. The platform is designed for analysts or other data consumers and is suitable for use as a testbed for evaluating emerging capabilities in a real-world environment.
A microservice framework to pull ML out of the lab and into production.

DATA LAYER
The WOPR data layer is a thin but innovative layer built on top of S3-compatible object storage for high scalability. This layer supplies data to the WOPR end-user environment, making data more accessible to end users while shortening the delivery time. It also maintains full provenance from enriched, translated records back to original source records or documents.

DATA PIPELINE
The WOPR data pipeline is scaled up to ingest and enrich large volumes of data, with a throughput of 250,000 records per minute. A configuration-based data translation component, including a visual workbench, accelerates data engineering tasks. Key components are containerized and tested under Kubernetes. The pipeline provides a turnkey extract, transform, load (ETL) solution for hydrating the WOPR data layer with disparate datasets. It can also be used independently to directly target Elasticsearch and similar relational database management systems (RDBMSs).

ANALYST ENVIRONMENT
The WOPR analyst environment is a modern, broad application enabling rapid implementation of the latest ML advances. WOPR’s search providers are pluggable. Five are included and Elasticsearch is supported out of the box. The recommendation capability pushes new and relevant content to analysts. Analysts collaborate and curate data into playlists, effectively labeling the data for future ML capabilities. ML as a service (MLaaS) is seamlessly incorporated to allow information extraction from documents and playlists.

SUPPORTING SERVICES
The WOPR environment is composed of several containerized microservices that can be used within the end-user environment, data pipeline, or independently, enabling WOPR to be applied to a variety of missions. Individual microservices can also be used outside the WOPR environment.

**Summarizer**
Generate a summary of a document or playlist.

**Elastomeric**
Perform large list queries against Elasticsearch.

**Vault**
Provide RESTful access to data in the WOPR playlists.

**Meridian**
Generate artifacts such as map layers from WOPR playlists, timelines, and entity manifests.

**Similarity**
Handle similarity searches and content recommendations.

**AdaptNLP-based Services**
Deliver entity extraction and a question and answer feature.