**Capability Case: Automated Content Tagger**

**Intent:**

To provide a way for people to add annotations to electronic content. Annotations can be comments, notes, explanations and semantic tags. Content annotation provides metadata that is becoming increasingly important to improve the precision of search as well as context-based information retrieval and repurposing.

**Solution Story:**

**AeroSWARM Automated Markup, HighWire Stanford e-Library**

The creation of markup from unstructured text sources such as web pages is tedious and time-consuming. Anyone who produces documents on a regular basis (e.g., intelligence analysts, commanders), or who has a large quantity of legacy documents, needs some form of automated markup assistance. Lockheed Martin has built a tool called AeroSWARM, which reduces the effort required for markup. It automatically generates OWL markup for a number of common domain-independent classes and properties. The author can then manually do markup additions and corrections to the output of AeroSWARM.

A user can specify the set of web pages to mark up, and choose a target ontology. Then, AeroSWARM generates OWL markup like that shown in the figure above. The sample markup includes entities (e.g., person, place, organization), relations (e.g., Pinochet persToLoc Santiago) and co-references (e.g., Pinochet sameIndividualAs Augusto Pinochet). A table on the AeroSWARM site describes all the entities and relations that can be automatically identified and marked-up.

**Benefits:**

Reducing the workload for annotated markup to a level manageable by human effort. The markup enables valuable business capabilities such as improved search, information discovery and retrieval.

**Featured Products:**

AeroSWARM from Lockheed Martin

**Applicable Products:**

ClearTags from ClearForest, Enigma 3C Platform from Enigma, Semio Tagger from Semio, SKIP from SemanTx Life Sciences, Stratify Discovery System from Stratify