

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 meta-data 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**

Chile's Pinochet Hospitalized After Fall AP Associated Press
26 minutes ago

SANTIAGO, Chile - Former dictator Gen. Augusto Pinochet (news web sites) was hospitalized Friday after falling and breaking his shoulder. A close associate to Pinochet, retired Gen. Luis Cortez, said Pinochet fell at his countryside residence near Santiago and was rushed to the hospital by helicopter. Santiago Military Hospital officials said the 88-year-old former ruler was to be released later Friday. Pinochet has retired from public life; associates say he devotes most of his time dealing with his poor health. Last November, Pinochet was admitted to the hospital with a broken left wrist after falling at home. He has a pacemaker, suffers from diabetes and arthritis and has sustained at least three minor strokes since 1998. His health problems led Chilean courts to drop charges filed against him for alleged human rights abuses during his 1973-90 regime. Some 3,197 people were killed for political reasons during his long reign, according to a report by the civilian government that succeeded Pinochet in 1990.

```

<?xml version="1.0"?>
<rdf:RDF
  xmlns:cm="http://ubot.lockheedmartin.com/ubot/2003/10/aerocSWARM-ont.owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <c:Person rdf:ID="Pinochet">
    rdfs:label="Pinochet">
    <owl:sameIndividualAs rdfs:resource="#AugustoPinochet">
      rdf:type="http://localhost:3100/aerocswarm/2003/10/aerocSWARM-ont.owl#Person">
      rdfs:label="Augusto Pinochet"/>
    <c:persToLoc rdf:resource="#Santiago">
      rdf:type="http://localhost:3100/aerocswarm/2003/10/aerocSWARM-ont.owl#Place">
      rdfs:label="Santiago"/>
    </c:Person>
    <c:Place rdf:ID="SANTIAGOChile">
      <rdfs:label>SANTIAGO, Chile</rdfs:label>
    </c:Place>
    <c:Organization rdf:ID="SantiagoMilitaryHospital">
      <rdfs:label>Santiago Military Hospital</rdfs:label>
    </c:Organization>
  </rdf:RDF>
  
```

AeroSWARM finds references in a news story, and renders them as semantic mark-up

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