



Proudly Operated
by **Battelle** Since 1965

Battelle Number(s):

13938-E

Patent(s) Issued

Available for licensing in all fields

Available Technologies

Universal Parsing Agent™

SUMMARY

Information workers using available extraction tools often spend more time cleaning, sorting, and reformatting data in preparation for analysis than analyzing the data itself. The Universal Parsing Agent (UPA) is a document analysis and transformation program that helps mitigate this problem by supporting massive-scale conversion of information into forms suitable for the semantic web. UPA provides reusable tools to analyze text documents, identify and extract important information elements, enhance text with semantically descriptive tags, and output the information that is needed in the format and structure that is needed.

ADVANTAGES

- * Outputs information in XML format, which is easily usable by most popular database programs, including Microsoft® Access®, Microsoft® Excel®, Oracle®, IN-SPIRE™, and Starlight™
- * Identifies errors by updating status information as it parses data and provides a user-friendly way to respond to errors
- * Handles unexpected data content
- * Provides an easy-to-use testing environment that allows users to ensure they are picking up precisely the data they need
- * Creates unified data formats for research
- * Enhances raw information with new semantic tags; for example, generating results in human- and machine-readable XML format
- * Prepares historical data for the Internet (for example, converting data from ASCII to XML for web presentations)
- * Parses email survey responses
- * Parses third-party online calendars
- * Conducts competitive intelligence analysis.



Patents & Intellectual Property

- » IP Summary

Technology Portfolio(s)

- » Information Analytics and Visualization

Potential Industry Applications

- » Aerospace & Defense
- » Agriculture & Mining
- » Automotive & Transportation
- » Chemicals
- » Communications & Media
- » Computers & Electronics
- » Consumer Products
- » Energy & Utilities
- » Entertainment & Recreation
- » Fabric & Apparel
- » Food, Beverage & Tobacco
- » Healthcare, Pharma, Biotech & Medical
- » Manufacturing & Warehousing
- » Oil & Gas
- » Public Administration & Government
- » Recycling & Waste Management
- » Security
- » Wood, Paper & Forestry

John T. McEntire
Pacific Northwest National Laboratory
(509) 372-6960
john.mcentire@pnnl.gov
<http://availabletechnologies.pnnl.gov>



Proudly Operated by **Battelle** Since 1965