

Data Sheet

Interwoven® MetaTagger™ 4.1

Overview

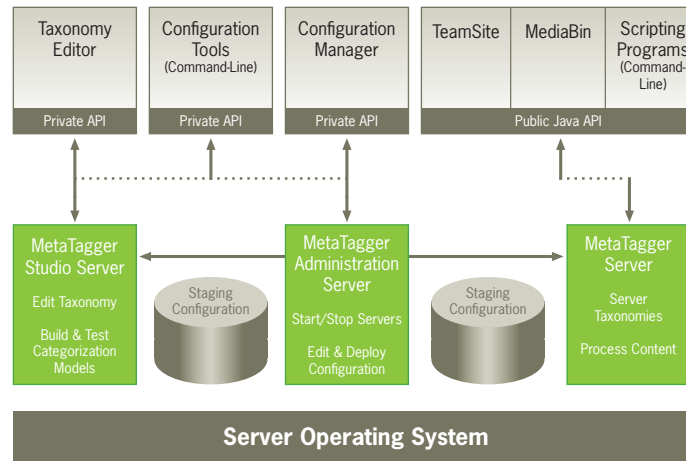
Interwoven MetaTagger™ Content Intelligence Server helps enterprises accelerate business processes, improve message impact, reduce business risk and lower operational costs by consistently and efficiently organizing unstructured information that drives their business—including office documents, email, and web content.

Based on business requirements

MetaTagger intelligently, automatically categorizes content and extracts information, based on business requirements, organizational standards, taxonomies, and collective knowledge. Supporting interactive, guided and fully automated operation, MetaTagger minimizes tedious, time-consuming, labor-intensive, and inconsistent manual tagging. Using MetaTagger, you can be sure the metadata you rely on for content presentation, navigation, routing, and search are consistent, accurate and up-to-date.

Modular design

A key component of Interwoven ECM, MetaTagger provides advanced taxonomy, metadata and content intelligence services as part of an Interwoven TeamSite® Content Management implementation. MetaTagger can also be used separately to provide these services to composite applications for content routing, filing, auditing, and presentation.



Product architecture

The Interwoven MetaTagger Content Intelligence Server provides an open suite of network services and clients for taxonomy development and navigation, content categorization and metadata extraction. Open, flexible, and scalable, MetaTagger's architecture supports integration with any content repository and any application—to process and deliver content across the enterprise.

At the core of each MetaTagger implementation are three servers:

- **MetaTagger Server** provides taxonomy navigation, categorization and metadata extraction functions
- **MetaTagger Studio** server provides taxonomy development functions,

- **MetaTagger Administration Server** provides configuration and service management support for the other two servers

Complementing these three servers, is the MetaTagger collection of graphical and command-line clients for taxonomy development and system configuration. MetaTagger's Java and command-line scripting API's deliver integration with composite, or Web, applications.

Dynamic content delivery

Dynamic content delivery lets portal developers place content based on attributes such as audience, product, or document type.

Knowledge directory

Knowledge directories organize—by taxonomy—a large collection of content resources, so that it's easier for content developers to find what they need

Intelligent content routing

Intelligent content routing uses business processes to distribute inbound requests and content to the right people automatically determining tacit attributes, such as topic and document type.

Content/metadata audit

Content/metadata auditing utilizes automatic categorization and information extraction to verify accuracy and consistency of key metadata required to support business processes.

Interwoven is a global leader in content management solutions

Interwoven's software and services enable organizations to effectively leverage content to drive business growth by improving the customer experience, increasing collaboration, and streamlining business processes in dynamic environments.

Our unique approach combines user-friendly simplicity with robust IT performance and scalability to unlock the value of content.

Today, nearly 3,800 enterprise and professional services organizations worldwide have chosen Interwoven, including: adidas, Airbus, Avaya, Cisco, DLA Piper, the Federal Reserve Bank, FedEx, HSBC, LexisNexis, Microsoft, Samsung, Shell, Samsonite, White & Case, and Yamaha. Over 18,000 developers and over 300 partners enrich and extend Interwoven's offerings. To learn more about Interwoven, please visit www.interwoven.com.

Interwoven, Inc.
803 11th Avenue
Sunnyvale, CA 94089 USA
(408) 774-2000

Key Features

Key feature	What it does
Automatic categorization	Automatically identifies the right categories for a document in one or more taxonomies; MetaTagger supports both contextual recognition, categorizing content based on key words, phrases and contextual cues, and categorization-by-example, categorizing content by comparison to expertly classified training documents
Summarizing and keyword extraction	Automatically generates content summaries that can vary in size from a few words to several sentences in length suitable for portal extracts
Entity extraction	Automatically pulls names, places, dates and other key information, using an extensible pattern recognition engine and compiled knowledge bases
Content profiling and taxonomy generation	Examines existing collections of content to automatically suggest a taxonomy or list of key concepts to assist in developing taxonomies
Document structure analysis	Automatically extracts key sections, such as title and date, from content, based on formatting and structural cues
Metadata validation and taxonomy mapping	Applies rules for combining and standardizing metadata, including adding and removing tags and converting between taxonomies
Graphical taxonomy editor	Provides a powerful visual development environment for building, editing, searching, refining, testing, and tuning custom or industry-standard taxonomies
XML-native format	Supports XML formats for taxonomy, validation rules, and entity extractor definition, content input and metadata output
Multiple format support	Includes built-in support for more than 150 file formats, including Microsoft™ Office, Adobe Acrobat™ and HTML
Extensible architecture	Enables solution developers to integrate additional content analysis and format support engines seamlessly into the MetaTagger pipeline