



## Inxight Software Government Solutions

11951 Freedom Drive, Suite 1300, Reston, VA 20190

Phone: 703.251.4429 | Fax: 703.251.4440

www.inxight.com | sales@inxight.com



# Inxight Entity Extraction for Counterterrorism and Intelligence

Inxight Software's counterterrorism solutions help analysts and field personnel automatically identify, extract and classify mission-critical entities (people, places, addresses, dates, organizations, weapons, vehicles, etc.) in unstructured text in multiple languages, providing faster, more accurate actionable intelligence. Entity extraction is an essential tool for operations that require link analysis, event tracking and order of battle analysis.

**Inxight ThingFinder®** is a powerful text-analysis application that automatically identifies and tags more than 25 entity types in a document. ThingFinder can also be trained to recognize list-based entities such as people of interest.

**Inxight ThingFinder Advanced** extends these capabilities by allowing users to define custom entities and simple link types based on patterns in text, leveraging Inxight's sophisticated language analysis as the starting point for pattern-building. New entity types for the counterterrorism effort are now available, including weapons, facilities, vehicles, geocoordinates and MGRS numbers.

Analysts currently have to "hand-chunk" text manually to load meaningful elements (names, places, organizations, geocoordinates, etc.) into databases for storage and search. Inxight ThingFinder instantly automates this hand-chunking process: There's no need for highlighters on hard copy, and no more hand-loading of data into large repositories. Inxight quickly auto-loads this information into databases, where analysts can then perform their queries and analysis, instead of spending time loading data.

Streaming text, message traffic, cables, reports – the data is voluminous. And analysts don't always know which piece of text needs to be connected to another piece. Any of this data

could be the next terrorist plot, the next bomb plot, the next dirty bomb, the next assassination attempt, etc. Agencies can't afford not to automate this process.

## Collection - Analysis - Production - Dissemination

Inxight helps with The Intelligence Cycle by picking up where collection leaves off. Vast quantities of textual data stream in and need to be categorized, summarized and clustered, and mission-critical data elements must be extracted, stored appropriately, and linked within and across documents.

Inxight does all of this in a single platform. Whether it's analysis or production, analysts will benefit from a tool that is a turnkey solution to the "text problem."

Inxight ThingFinder and ThingFinder Advanced enable you to:

- Discover and extract more than 25 entity types in unstructured text, even those not yet on watch lists.
- Extract entities in more than 20 languages, including Arabic, Chinese, Farsi, Korean and Spanish.
- Perform aliasing, variant detection, and relevance ranking.
- Monitor watch lists using the Name Catalog.
- Define custom entities and link extraction.

U.S. government intelligence agencies are deploying ThingFinder in operations where information is the battlefield and the fight is to quickly find critical information buried in mountains of intelligence, to locate terrorists, their whereabouts, determine their actions, associations and future plans.

*"We see Inxight as a key strategic partner, as there is unquestionable value in how their technology, particularly their language capabilities, can help government and commercial organizations identify important information. A critical factor in our decision to work with Inxight is that it has already proven that its solutions deliver 'high fidelity' results in very demanding information environments."*

- Gilman Louie  
CEO, In-Q-Tel



## Features

Inxight's true understanding of natural language – language-aware tokenization, part-of-speech tagging and noun phrase identification – powers the automatic extraction and classification of entities in unstructured text.

### Inxight ThingFinder Example

DOCUMENT 1: (UNCLASSIFIED) FAWIZ AL ((RABBATI)) IS RECRUITING PEOPLE IN THE AL-KRKH DISTRICT, MOSUL TO CONDUCT COORDINATED TERRORIST ATTACKS AGAINST COALITION FORCES FOR \$500. THE MAIN WEAPONS USED IN THESE ATTACKS ARE ROCKET PROPELLED GRENADES, AND AK-47S. RABBATI IS SIXTY-SEVEN INCHES TALL, AND WEIGHS 185 TO 190 POUNDS. HE HAS A MUSCULAR BUILD AND VERY DARK SKIN. RABBATI OWNS A RESIDENTIAL BUILDING ON SHEIKH FAARQUM STREET. RABBATI RECRUITED JABIL (NAATAN) TO WORK FOR HIM IN 1998, NFI.	
Person	Fawiz al Rabbati, Rabbati Jabil Naatan
Place	Al-Krkh District, Mosul
Currency	500 USD
Year	1998
Weapons	rocket propelled grenades AK-47
Address	Sheikh Faarqum Street
Noun Group	coordinated terrorist attacks coalition forces residential building
Measure	sixty-seven inches 185-190 pounds

### Variant Identification and Grouping

Variant identification and grouping allow ThingFinder to accurately classify all relevant entities in a document, even one-word entities, and to provide true counts reflecting the number and location of ALL appearances of a given entity.

### Normalization

Entities such as dates, currencies and measurements are normalized into standard formats, taking the guesswork out of the metadata creation, data mining and link analysis.

### Relevance Ranking

The entities extracted by Inxight ThingFinder are given relevance scores, reflecting their importance to the document as a whole, making ThingFinder an essential part of any data categorization solution.

### Customization

ThingFinder can be easily extended to identify and extract custom list-based entity types, such as names of persons of interest. ThingFinder Advanced enables you to define custom entity and link types for pattern-based discovery and extraction in 31 languages.

### Other Inxight Products for Counterterrorism

**Inxight StarTree®** is a visualization tool used by many agencies to reveal and explore relationships.

**Inxight SmartDiscovery Awareness Server™** is a powerful federated search and alerting solution that can provide access to hundreds of open source and classified sources from a single interface.

See details at [www.inxight.com](http://www.inxight.com).

## About Inxight

Inxight Software, Inc. is the leading provider of enterprise software solutions for information discovery. Using Inxight solutions, companies can better discover, retrieve and connect with unstructured, semi-structured or structured text. Inxight is the only company that provides a complete, scalable solution to enable information discovery in all major languages. Customers include enterprise companies such as Air Products, Novartis, Procter & Gamble and Thomson, multiple government agencies, including the Department of Defense, Defense Intelligence Agency and Department of Homeland Security, and software OEMs such as SAP, SAS, IBM and Oracle. Inxight is an In-Q-Tel portfolio company.

Copyright ©2005 Inxight Software, Inc. All rights reserved. Inxight, StarTree, SmartDiscovery, ThingFinder and Awareness Server are trademarks or registered trademarks of Inxight Software, Inc. All other products are trademarks of their respective companies.

## Technical Specifications

Inxight ThingFinder and ThingFinder Advanced are available as part of Inxight SmartDiscovery® 4.2.

- Single API for easy integration into enterprise systems and applications; hooks for integration are accessible via C++, Java and SOAP. Data is transferred via XML.

## Operating Systems

- Microsoft Windows 2000, XP and 2003 Server
- Sun Solaris 8.0 and 9.0
- RedHat Linux Enterprise Server 2.1 and 3.0.

Please contact Inxight for other platforms.

## Language Support

Arabic, Bokmål, Danish, Dutch, English, English-MTF, Farsi, Finnish, French, German, Italian, Japanese, Korean, Nynorsk, Portuguese, Russian, Simplified Chinese, Spanish, Swedish, and Traditional Chinese.

## Entity Type Support

Below are some of the major entity types that ThingFinder supports out of the box\*:

- Person
- Place
- Address
- City
- Country
- State
- Internet Address
- Vehicles
- Weapons
- Facilities
- Company
- Organization
- Numeric Entities
- Date and Time
- Noun Group, Proper Nouns

\*NOTE: Not all entity types are supported in every language.