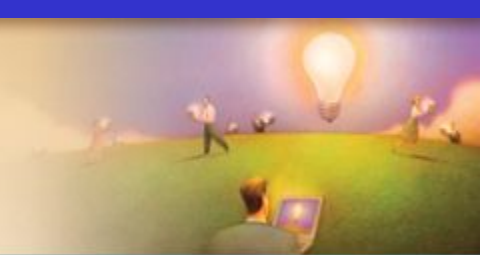


Innovations by InterSystems



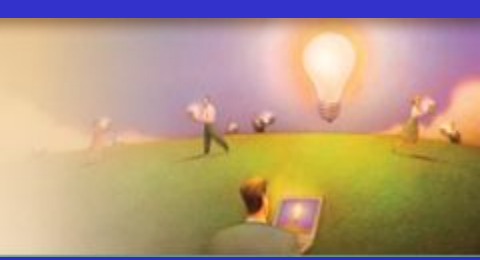
iKnow and DeepSee

Agenda



- What is iKnow?
- Semantic Analysis.
- %iKnow.Queries
- Matching Dictionaries.
- %iKnow.Semantics.
- Newer features:
 - Attribute Customizations.
 - iFind.
 - Text Categorization.
- iKnow features in DeepSee.
- Configuring iKnow and DeepSee.

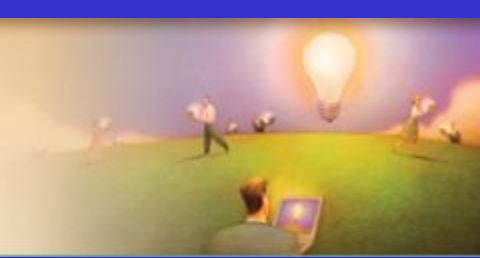
What is iKnow?



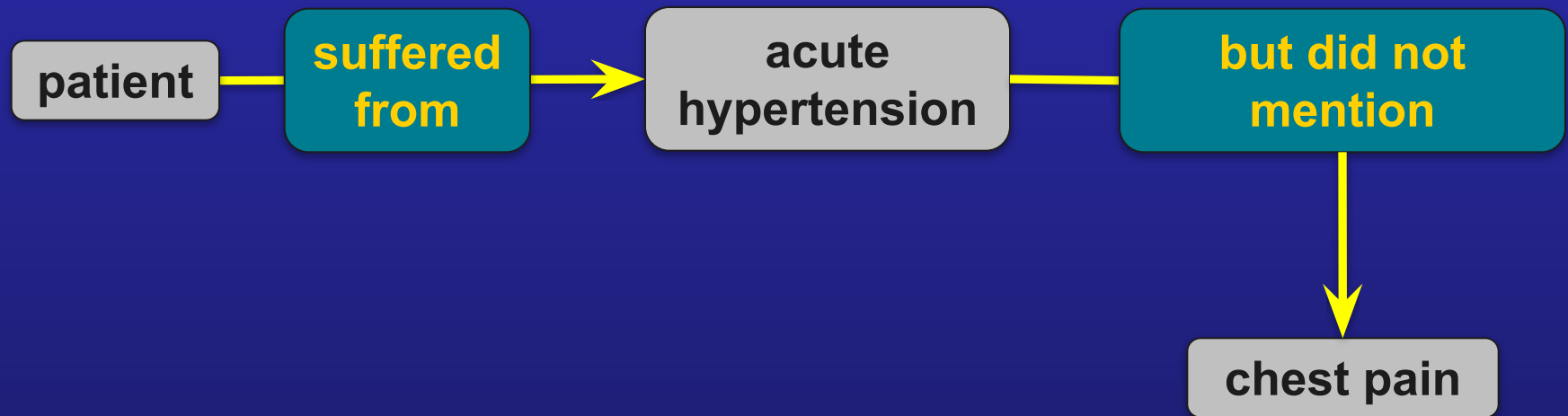
- iKnow is a semantic analysis tool.
 - Indexes the *concepts* and *relations* within text for querying and analysis.
 - Uses language models rather than training data or ontologies to detect relations.
 - Supported languages: Dutch, English, French, German, Portuguese, Russian, Spanish, Ukrainian, Swedish*, and Japanese*.
 - Sources of text include: Plain text files, SQL fields, social media.

*Support added in 2016.1 release.

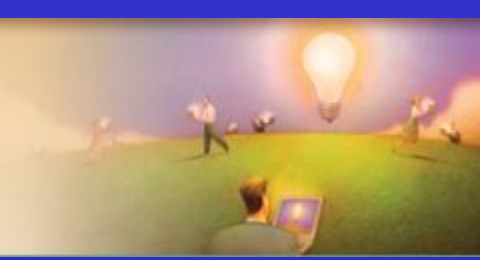
Semantic Analysis: Relations, Concepts, Negation



- The patient suffered from acute hypertension but did not mention any chest pain.



Semantic Analysis Results



Concept

patient

acute hypertension

chest pain

Relation

suffered from

but did not
mention

Concept-Relation-Concept

patient *suffered from* acute hypertension

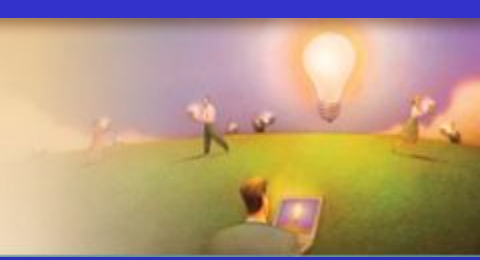
acute hypertension *but did not mention*
chest pain

Importance of Language Models



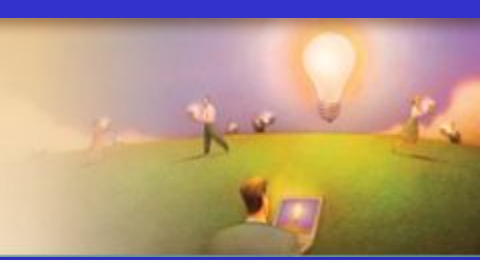
- iKnow indexing is subject matter neutral.
 - A language model applies to any text written in the language: medical, legal, scientific, business, and so on.
- iKnow indexing automatically detects meaningful word groups.
 - Labels “acute hypertension” and “chest pain” as concepts.
 - Labels “but did not mention chest pain” as a negation context.
- No need for ontologies or training data.

%iKnow.Queries



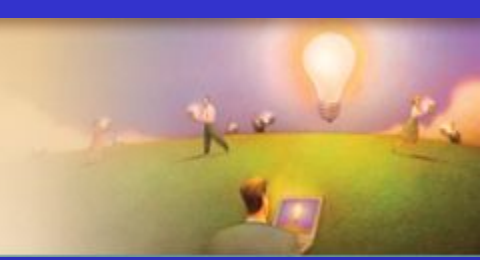
- Includes:
 - GetTop() – Most frequently occurring entities across a set of sources.
 - GetRelated() – Entities in a relationship with the supplied entity.
 - GetByEntities() – All CRCs or paths containing a particular set of entities.
 - GetSummary() – Most relevant sentences in a source.
 - GetSimilar() – Entities similar to a given entity.

Matching Dictionary




- User provided group of related terms.
 - Provides external (domain) knowledge to iKnow results.
 - Allows for coarser grained analysis.
 - Example (2001 A Space Odyssey):
 - hal □ hal.
 - hal9000 □ hal.
 - heuristic algorithm computer □ hal.
- iKnow smart matching mechanism returns a match score.
 - Configurable threshold for matches.

iKnow Architect (2016.1)



- Management Portal Tool for creating, configuring, and managing iKnow domains.
 - Domain Settings, Metadata, Data Locations, Blacklists
 - Compile and build domains.
 - Launch indexing and knowledge portal pages.
- Some iKnow features not supported by Architect. Edit class definition using IDE.
 - Matching Dictionaries.

Demonstration



<< >> movies ▼

sentence: to

This is a sample User Interface built on top of the iKnow APIs, visualizing the results of the iKnow Smart Indexing API. You can either select an existing source from the dropdown list, or use the input button to enter free text directly.

[hide](#)

Indexed sentences

The opening crawl reveals that the Galactic Empire has pursued the Rebel Alliance across the galaxy, forcing them to establish a secret base on the remote ice planet Hoth.

Dark Lord of the Sith Darth Vader (David Prowse/voice: James Earl Jones) sends robotic probes in search of the base and its commander, Luke Skywalker (Mark Hamill).

While Luke is patrolling near the base, he's knocked unconscious by an indigenous predator, the Wampa.

Back at the base, the smuggler-pilot Han Solo (Harrison Ford) announces his intention to leave the Rebels and pay the debt he owes to the gangster Jabba the Hutt, much to the displeasure of Princess Leia (Carrie Fisher).

After Han discovers that Luke has not returned from patrol, he delays his departure and leaves the base to search for him.

Luke escapes the Wampa's lair but is overcome by the cold.

He sees an apparition of his late mentor, Jedi Master Obi-Wan Kenobi (Alec Guinness), who instructs him to receive training from Jedi Master Yoda (voice: Frank Oz) on the planet Dagobah.

Han finds Luke and gives him shelter until they're rescued the following morning.

When ground scans detect an object outside the base perimeter, Han and Chewy investigate, and find an Imperial probe droid that transmits the location of the Rebel base to the Imperial fleet before firing upon Chewy and being destroyed.

On the command ship of Darth Vader the droid's transmission is picked up by Captain Piett, but his superior, Admiral Ozzel, dismisses the message until Vader sees the transmission and realizes it is the Rebel base.

He overrules Ozzel and orders the Fleet to the Hoth system.

Later, as the Fleet emerges from trans-star warp, Vader is informed by his infantry commander, General Viers, that the Rebels have set up infantry trenches and an energy shield to protect them from the Empire's orbital bombardment.

Ordering Viers to launch a surface attack, Vader is furious that Ozzel has clumsily given away surprise, and he communicates this displeasure to Captain Piett as he promotes him on the spot to Admiral while Ozzel is given Vader's own fatal brand of reprimand.

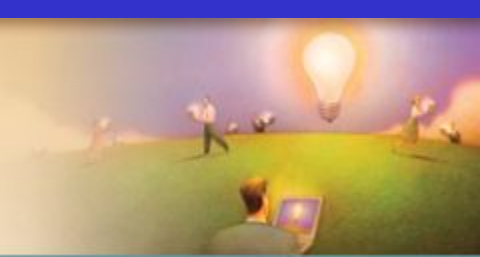
The Imperial forces land their ground assault walkers beyond the energy shield and Luke leads his squadron of flying speeders into battle.

%iKnow.Semantics (2012.2+)



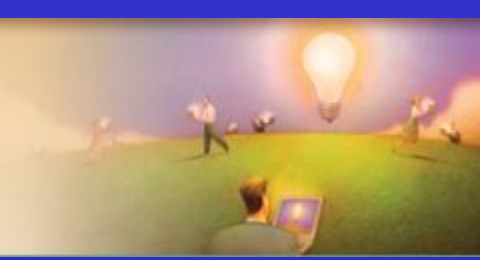
- Introduces concept of *dominant* entities.
 - Most important entities not most common.
 - Algorithm revised for 2015.2 release.
 - Explained in documentation.
- Includes queries:
 - GetBySource() – Dominant elements in a specific source.
 - BuildOverlap() – Generates dominant term overlap information for all sources in a domain.
 - FindMostTypicalSources() – Most typical sources.
 - FindBreakingSources() – Most atypical sources.

Attribute Customizations



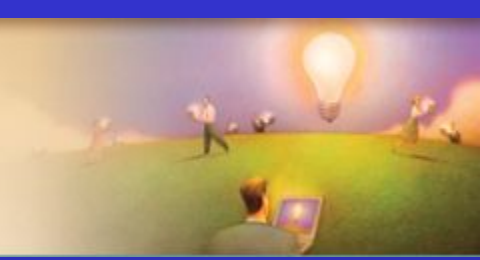
- Negation.
 - Augment default markers with additional markers for particular use cases.
- Sentiment.
 - No default markers.
 - Supply custom sentiment markers.
- Attribute markers.
 - Supply custom markers in User Dictionary.
 - iKnow performs attribute tagging during loading.

iFind



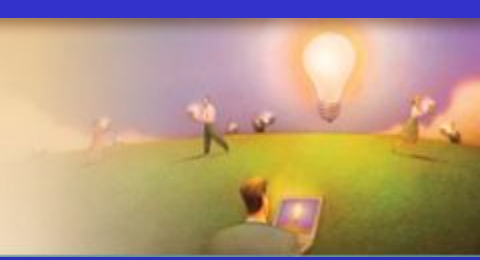
- SQL feature for performing text search.
 - Add iFind index to columns containing text.
 - Include iFind index syntax in WHERE clauses of SQL queries.
- Support for the following searches:
 - Stemming and de-compounding.
 - Word and word phrase search.
 - iKnow entity search.
 - iKnow semantic search using path, proximity, and dominance information.

Text Categorization



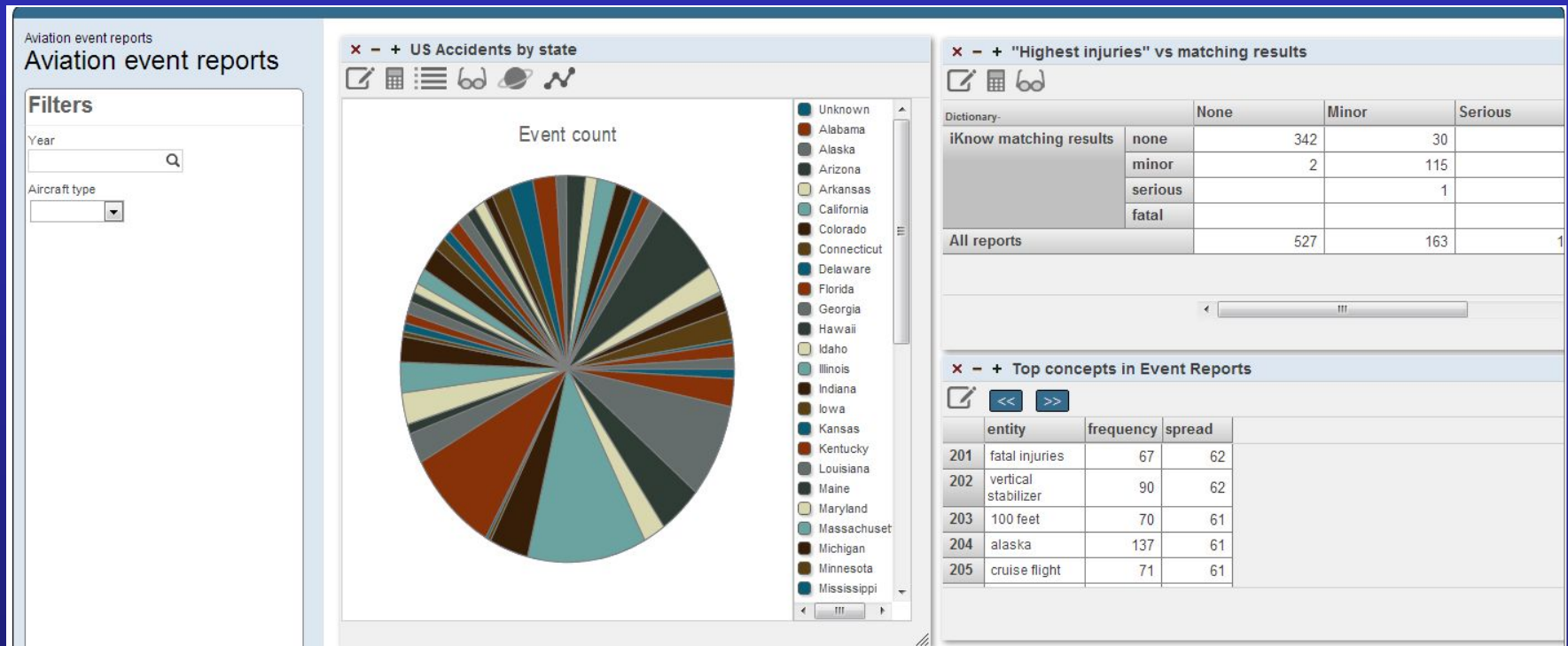
- Label (categorize) source texts based on their contents (entities and relations).
 - Create a classifier by analyzing an existing (training) set of already labelled texts
 - Apply classifier to new and as yet unlabelled texts.
- Wizards available for building and testing classifiers.
 - System Explorer □ iKnow □ Text Categorization

DeepSee and iKnow

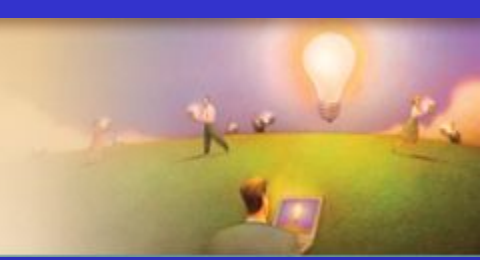


- DeepSee cubes can include iKnow indexing results and analyses:
 - iKnow Dimensions.
 - Entities (concepts and relations).
 - Dictionary matching results.
 - Use as rows, columns, and filters on pivot tables just like data and time dimensions.
 - Detail Listings.
 - iKnow summaries.
 - Content Analysis Plugin to allow users to perform a variety of iKnow analyses on text sources.

Demonstration

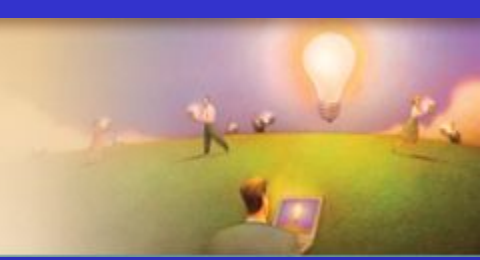


iKnow Dimensions



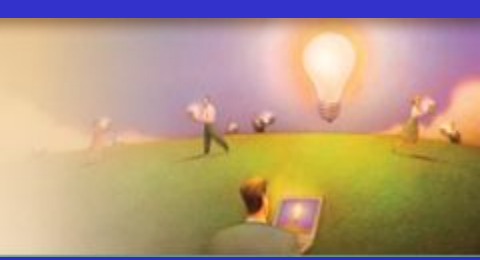
- Entity dimension.
 - Single level.
 - Members are entities (concepts or relations).
 - Analyzer displays first 100 in decreasing order by spread.
 - Filter options contain all entities. Searchable.
- Dictionary dimension.
 - Level 1: one member for each dictionary.
 - Level 2: one member for each item containing all matches for that item.
 - Matching dictionaries loaded as termlists.


iKnow Measure



- Connects unstructured data to cube.
 - Purely configuration. Not visible to Analyzer.
 - Connects DeepSee cube to text sources and dictionaries.
- Referenced by iKnow dimensions.

Content Analysis Plugin



- Launch from Analyzer or Dashboard.
 - Select cell and click 
- iKnow features include:
 - Content Analysis.
 - Typical and breaking sources.
 - Entity Analysis.
 - Overview: frequency and spread for 10 most common groups.
 - Cell breakdown: distribution of entities selected on Overview tab.
 - Entities: frequency and spread for entities similar to entity selected on Cell breakdown.

Demonstration

Menu Home | About | Help | Logout DeepSee > Architect

AviationEvents Server: DENELSONE6410 Namespace: SAMPLES Switch
User: UnknownUser Licensed to: ISC Learning Services - Instructors Instance: CACHE

DeepSee by InterSystems

View: [Icon] New Open Save Compile Build Documentation

Architect

Source Class Model Elements Add Element Undo Expand All Collapse All Reorder

Aviation Event

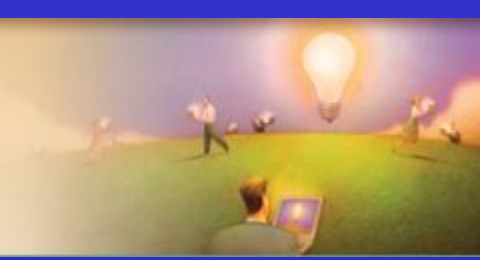
- %ID
- Aircraft1
- Aircraft2
- AirportDirection
- AirportDistance
- AirportElevation
- AirportId
- AirportLocation
- AirportName
- Altimeter
- City
- CoordsLatitude
- CoordsLongitude
- CoordsSource
- Country
- CountryCode
- DayOfWeek
- EventDate
- EventId
- FAADistrictOffice
- Genmetar
- InjuriesGroundFatal
- InjuriesGroundMinor
- InjuriesGroundSerious
- InjuriesHighest
- InjuriesTotal
- InjuriesTotalFatal
- InjuriesTotalMinor
- InjuriesTotalNone

AviationEvents	Element Type	Details
Measures		
SeatsTotal	number measure	Aircraft1.SeatsTotal
InjuriesTotal	number measure	InjuriesTotal
Report	iKnow measure	Narrative1.Report
Dimensions		
AirportDim		
H1	data dimension	
Hierarchy		
AirportName	level 1	AirportName
AirportLocationDim		
H1	data dimension	
Hierarchy		
AirportLocation	level 1	AirportLocation
LocationDim		
H1	data dimension	
Hierarchy		
Country	level 1	Country.Name
State	level 2	State.Name
City	level 3	City
EventDateDim		
H1	time dimension	EventDate
Hierarchy		
EventDateY	level 1	Year
EventDateM	level 2	MonthNumber
InjuriesHighestDim		
H1	data dimension	
Hierarchy		
InjuriesHighest	level 1	InjuriesHighest
LightConditionsDim		
H1	data dimension	
Hierarchy		

Details Tools

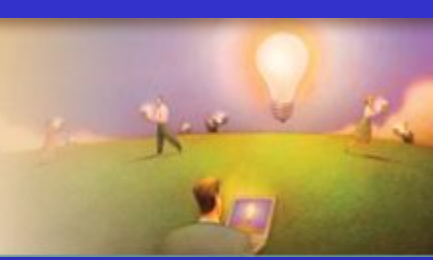
Select an item to view details

Configuring iKnow Measure



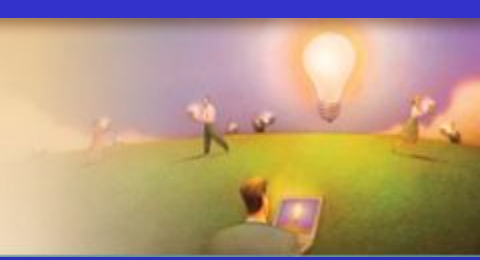
- iKnow Measure:
 - Source Values: Property or expression.
 - Aggregate: Count.
 - Type: iKnow.
 - iKnow Source: string, stream, file, or domain.
 - Dictionaries: select from available termlists.

Configuring iKnow Dimensions



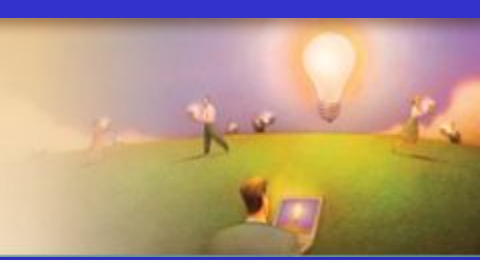
- Entity Dimension.
 - Dimension Type: iKnow.
 - iKnow Type: entity.
 - iKnow Measure: iKnow measure name.
- Dictionary Dimension
 - Dimension Type: iKnow.
 - iKnow Type: Dictionary.
 - iKnow measure: iKnow measure name.

iKnow Listing Features



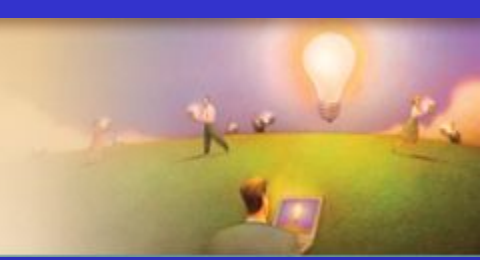
- Include iKnow summary.
 - \$\$\$IKSUMMARY[iKnowMeasure, summaryLength].
- Include content analysis plugin.
 - \$\$IKLINK[iKnowMeasure].
 - Allows users to see: summaries, dictionary matches, negation contexts, and dominant entities for selected source(s).

Suggested Reading



- Using iKnow.
- Advanced DeepSee Modeling Guide □
Using Unstructured Data in Cubes.

Summary



- What are the key points for this module?

