Article Benjamin De Boe · May 31, 2016 5m read

iKnow demo apps (part 1) - Knowledge Portal

InterSystems' iKnow technology allows you to identify the concepts in natural language texts and the relations that link them together. As that's still a fairly abstract definition, we produced <u>this video</u> to explain what that means in more detail. But when meeting with customers, what really counts is a compelling demonstration, on data that makes sense to them, so they understand the value in identifying these concepts over classic top-down approaches. That's why it's probably worth spending a few articles on some of the demo apps and tools we've built to work with iKnow.

In the first article in this series, we'll start with the Knowledge Portal, a simple query interface to explore the contents of your domain.

Installation and setup

This interface has shipped with the kit since the very first iKnow release and does not require any particular type of data, metadata or other external dependencies. It will work right off any domain in your system, and can be accessed from the link in the iKnow Architect page for your domain definition or by opening the following link: http://localhost:57772/csp/samples/iKnow.UI.KnowledgePortal.zen. If you don't have a domain set up yet, take a look at this article on creating a simple iKnow domain. Alternatively, you can initialize the demo domain on aviation event reports in the SAMPLES namespace by calling do

##class(Aviation.Utils).SetupStandalone(), which we'll use in this article.

Depending on your security settings, you may still have to enable your web application for iKnow (like for DeepSee) through the checkbox in the SMP page at System Administration > Security > Applications > Web applications for your web app or by calling do EnableIKnow^%SYS.cspServer("/csp/myapp/").

For 2016.3 (currently available in field test), we did a complete makeover of the original layout and styled it using the Bootstrap libraries available in the kit since 2016.1.

Using the Knowlege Portal

As this is a read-only interface for exploring the output of iKnow's bottom-up approach for a given domain, there is no particular workflow to run through. We'll stick to a simple investigative scenario, using the Aviation Events Demo domain in the SAMPLES namespace (see above). Upon opening the Knowlege Portal, you should see the following:

Explore		Explo
Top Concepts	frequency domin	nampe
airplane	6343	983
pilot	6200	1083
engine	2093	559
accident	1667	766
flight	1474	697
runway	1160	506
helicopter	933	138
examination	838	442
visual meteorological conditions	780	775
ground	708	431
		< >
Sources Paths CRCs		

Exploring entities

The widget in the top left displays the top concepts for this domain, which quite evidently contain aviation incident related terminology. The two numbers to the right of each entity are the number of times the entity occurs in total (frequency) and the number of distinct documents in which it appears (spread). If you toggle to "dominance" in the widget's title bar, the entities will be ranked according to dominance, a bottom-up metric expressing entity relevance within the domain. Using the "<" and ">" buttons at the bottom, you can scroll through the different pages of results.

When you click one of the entities in the list, the real exploration starts. Let's start with "pilot":

pilot		Explore!					1200 🔻	
Top Concepts	requency domin	ance	Similar Entities			Related Concepts	related prox	amit
airplane	6343	983	pilot	6205	1083	airplane	384	
tolic	6200	1083	private pilot	385	297	passenger	117	
ingine	2093	559	student pilot	319	141	autopsy	89	
ccident	1667	766	accident pilot	173	58	controller	112	
light	1474	697	commercial pilot	168	147	specimens	61	
unway	1160	506	private pilot certificate	128	117	engine	67	
elicopter	933	138	commercial pilot certificate	91	87	private pilot certificate	59	
xamination	838	442	certificated private pilot	78	76	helicopter	60	
isual meteorological conditions	780	775	pilots	75	58	takeoff	43	
ground	708	431	pilot's logbook	70	62	accident	51	
		< >			< >			<
Sources Paths CRCs								
1200 :SQL:2010:20101209X9403	36 The plot reduced th enough al engine po identified a fuel remai	sustained serior re engine power, rspeed or altitude wer on both enginand was related to ning, based upon	us injuries and his passenger sustain while turning to final for runway 17, to to reach the airport so he performed nes within the preceding few months o "fuel delivery that exceeds that req the information loaded by the plat	ted minor injurie when the left en d an off-field lan The piot f uired for stable prior to the filg	es In a gine lost ding ¹ felt that in combust jht	a written statement submitted by the pilot power The pilot reported that there The pilot wrote that he had encountere h both cases the root cause of the proble ion." The fuel flow indicator recorded (, he had just was not d a loss of m was never 36.8 gallons of	
1199 :SQL:2010:20101027X8553	34 The plot with Title 1	elected to ditch 14 Code of Feder	in the Atlantic Ocean about 2.6 naut al Reculations Part 91. and the loter	ical miles east of was not inium	of the airp ed Th	cort The personal flight was operated	in accordance 🛛 👁	

Now many more widgets display data.

- The similar entities widget shows all other entities iKnow found that are "similar" to the string or entity we were exploring, with similar defined as containing that string somewhere at the start of a word within the similar entity (note that for languages like German and Japanese, there's also an alternative option looking at any position inside the entities, cf the <u>\$\$\$IKPENABLENGRAMS</u> domain parameter). The frequency and spread numbers have the same meaning as in the top concepts widget.
- The related concepts widget shows all other concepts the entity being explored is directly related to (a single relational element between). The frequency and spread numbers now express how often the concepts appear in a direct relationship. If you toggle to "proximity" in the title bar, the relationship can now also be an indirect one, with the concepts appearing at a longer distance within the same path.
- The full-width widget below will show the sources (records), paths (sentence subsections) and CRCs (concept-relation-concept triples) in which the entity appears, with the seed entity highlighted in green. In the source view, text fragments containing the seed entity are displayed, with negative sections marked in red.

Rather than clicking an entity in any of these widgets, you can also start typing in the exploration box in the top left corner. The Knowledge Portal will start suggesting actual entities from the domain that match what you are typing so you can explore even faster.



For example, when starting from the term filter, the similar entities will include many different types of filters installed in aircraft, not because iKnow contained any predefined knowledge on aviation mechanics, but simply because it identified them from the reports bottom-up.

Exploring context

When looking at sources or paths in the main bottom widget, you can click the eye icon to consult the complete source text and optionally highlight indexing results or dictionaries (more on the latter in future posts). Using the controls at the bottom, you can also look up the source's metadata and generate summaries of the text. In other words, after initial exploration points you to a couple of documents, you can quickly find the original context and read a summary of the entire document.

iKnow demo apps (part 1) - Knowledge Portal Published on InterSystems Developer Community (https://community.intersystems.com)



Filtering

If the sources in your domain have relevant metadata associated, you can use that to zoom in onto a specific subset of records by clicking the filter icon in the top right corner. In the popup, you can create simple predicates or combine them using boolean logic. When you click apply, all exploration results will only display entities and figures for the filtered set of sources. The filer icon will also remain highlighted as long as any filter conditions are active and the number of sources in the domain is updated to reflect the filter conditions.

filer	Filters ×	Ŧ	٥
Top Concepts frequency	domina	proxim	RY.
pilot	442 Year BETWEEN 2005,2010 add filter hare 🗶 I magneto	1	1
airplane	405 AND - State - California add filler here 🗶	1	1
california	137 add filer here	1	1
engine	130	<	>
accident	96 metadata v Highestinjury v = v None add		
runway	79		
cfi	76 Clear Apply		
flight	67		
helicopter	57 10		
substantial damage	49 43		
Sources Paths CRCs			
1034 :SQL:2006:20060120X00103 There	were no contaminants found in the filter	۲	•
281 :SQL:2006:20060803X01081 An en	gine logbook entry dated July 10, 2006, noted that a mechanic adjusted the valve clearance, changed the oil and filter , and installed a new agneto, new spark plugs, new muffler, new crankcase oil seal, new vollage regulator, and an overhauled alternator	۲	•
left ma			

If your domain also includes blacklists (not in the demo domain), you'll see another icon next to the filter one to toggle which blacklists' entries should be suppressed from all displayed results. Finally, the cog icon still in the

same button bar will allow you to jump to other domains, to explore yet new horizons.

Next week, we'll zoom in on another iKnow demo app for building dictionaries off an existing domain, formalizing the bottom-up insights brought by iKnow into a structured knowledge base.

#InterSystems Natural Language Processing (NLP, iKnow)

Source URL: https://community.intersystems.com/post/iknow-demo-apps-part-1-knowledge-portal