

### From Documents to Knowledge: Chemistry Data for Real World Evidence

Jeffrey L. Nauss, PhD
Sr. Application Specialist, Linguamatics
ChemAxon User Group Meeting
San Diego, March 2019









## **About Linguamatics**



- Agile, scalable, real-time NLP-based text analytics
- Rapid relationship extraction and knowledge synthesis using unique and powerful blend of methods
- Commercial out-of-the-box, deep domain expertise

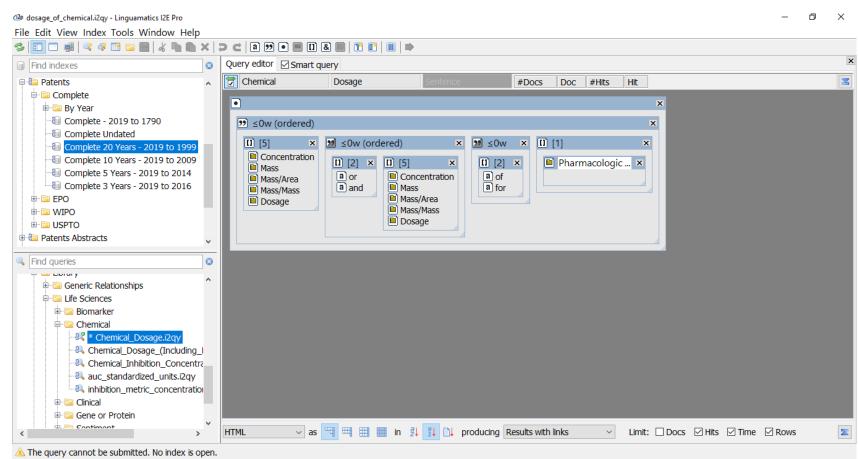
Pharma/Bioted	ch	Healthcare	Government			
Including 18 of	the top 20	Including major cancer centres, health systems, academic medical centres	Including FDA, NIH, NCI			
2	© 2019 Linguamatics		Linguamatics an IQVIA company			

### **Linguamatics I2E**



- Natural language processing (NLP)-based knowledge discovery platform
- Graphical view of the query
- Delivers results in real-time in structured tables
- Variety of other output formats
- Has an API
- Incorporates ChemAxon JChem











Chemical Relation Entity

> Arthritis,

inhibit > TNF

increase > TNF

Chemical

Cyclosporine inhibits

Rein

Gene Rein

Etanercept v dosage > 50 mg every 7d 52 > 23319853 5 > The aim of this study was to evaluate the proportion of patients with ankylosing spondylitis

CXCL8 are upregulated in

Disease

Psoriasis >

maintaining clinical remission after reduction of their subcutaneous etanercept dose to 50 mg every other week compared with that in patients receiving etanercept 50 mg weekly.

Factor-Alpha and Asymmetric Dimethylarginine Levels and Rescues Spatial Deficits in Young

Factor-Alpha and Asymmetric Dimethylarginine Levels and Rescues Spatial Deficits in Young

PPAR-α, ACOX, FAS, ACC and SCD protein levels, or increase of the LC3 II/LC3 I ratio and

However, neither GdCl3 nor etanercept could affect binge drinking-induced decrease of

37 > 29680792 2 > Pneumocystis jiroveci pneumonia in a patient taking Benepali for rheumatoid arthritis.

13 > 27445694 2 > Combined Intraperitoneal and Intrathecal Etanercept Reduce Increased Brain Tumor Necrosis

3 > 27445694 2 > Combined Intraperitoneal and Intrathecal Etanercept Reduce Increased Brain Tumor Necrosis

2 > 24346918 2 > Systemic sarcoidosis induced by etanercept: first Brazilian case report.

Rats after Bile Duct Ligation.

Rats after Bile Duct Ligation.

#Docs Doc

#Hits Hit

It was found out that CsA inhibits IL-8 production by stimulated THP-1 monocyte cell line in dose dependence course. There is ample evidence that several cytokines, including transforming growth factor-alpha (TGFalpha), interleukin (IL)-6, and IL-8 are upregulated in psoriasis, suggesting a pathogenic role for these

cytokines.



Assertion Clustering

Relationship Extraction

LONDON An advertisement that Pfizer placed in a French newspaper has stirred makers that the company might be positioning itself to play some sort of role Web in the contested, unsolicited bid by Sanofi-Synthélabo for Aventis.

day. Pfizer ran a full-page the newspaper, es Echos, that described the company's most popular pharmaceutical products with the tagline, "Health is like family, you build it."

It is common practice in France to take out battle. "Did they buy space cheap? Are advertisements that boast of a company's strengths before making a takeover bid. Merger specialists said the timing of the Pfizer ad, which comes as the legality of the Sanofi bid is being debated in a

"It has to be more than a coincidence," said Eric Cafritz, a partner with Fried Frank Harris Shriver Jacobson, who is based in Paris. "In these contested French bids, publicity is put to strategic use.

French courtroom, was puzzling.

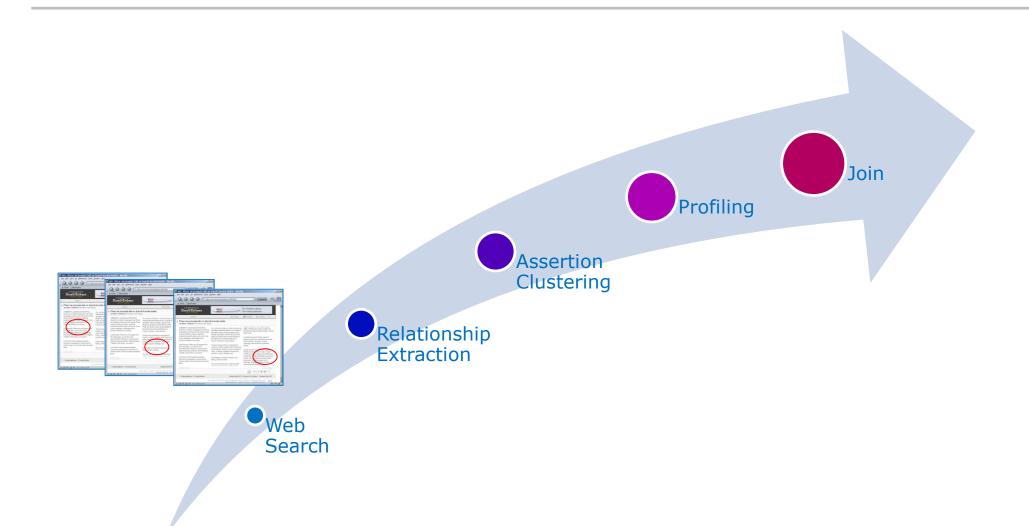
they trying to improve their image? We just

A spokeswoman for Pfizer, which is based in New York, said that the ad was in line with the company's overall message. She would not comment

Aventis, its larger rival on Jan. 26 and

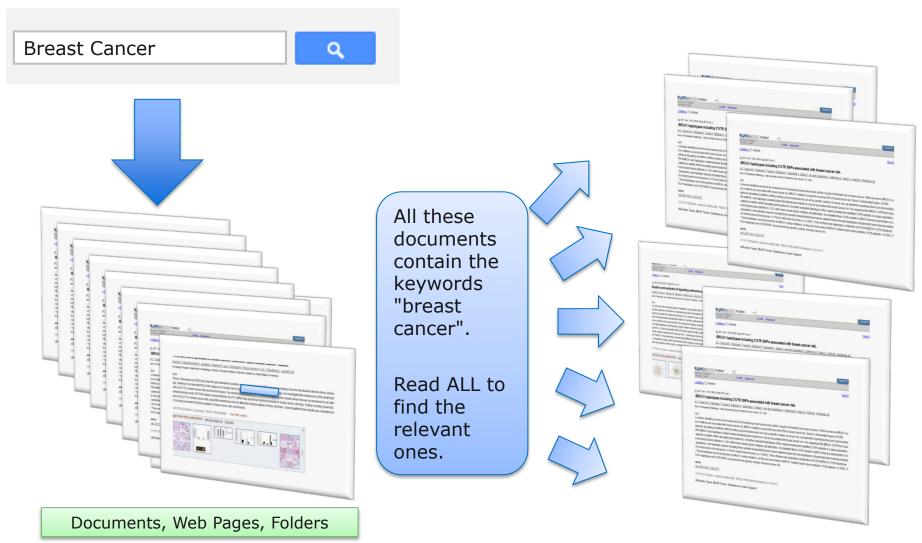


Search





#### **Web Search**

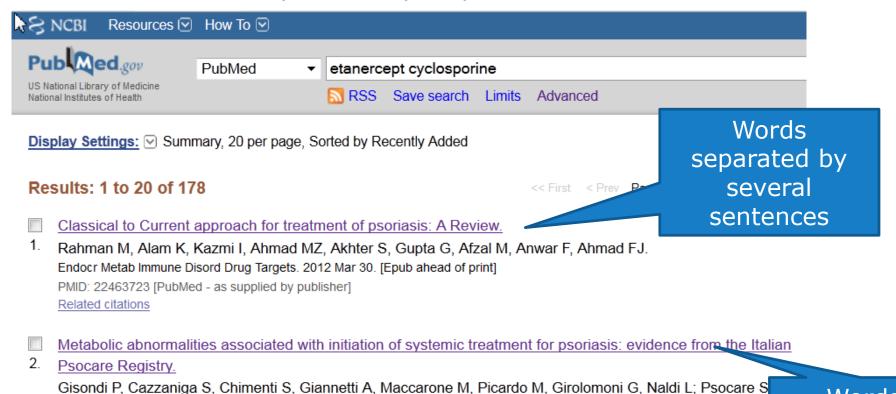




#### **Web Search of Two Terms**

Consider a search for "etanercept" and "cyclosporine"

PMID: 22313340 [PubMed - as supplied by publisher]



J Eur Acad Dermatol Venereol. 2012 Feb 7. doi: 10.1111/j.1468-3083.2012.04450.x. [Epub ahead of print]



Words in

adjoining

sentences

Related citations

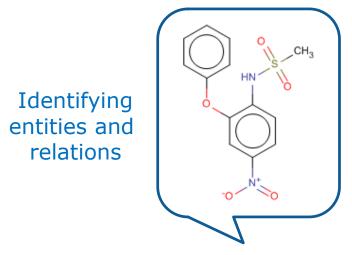
Group.

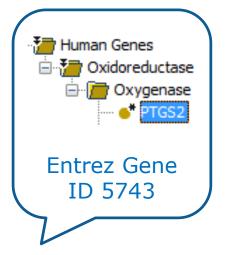
#### **Co-occurrence vs. Linguistics**

- Document co-occurrence is commonly reported in search engines results.
  - Great at showing there is some association, but not what the association is.
- Would a sentence co-occurrence be any better?
  - "Etanercept and cyclosporine were the drugs most frequently prescribed at entry."
  - "Etanercept demonstrated higher efficacy and good tolerability as replacement therapy for cyclosporine in plaque psoriasis."
- We can use linguistics to distinguish between the specific relationships



# **Linguistic Processing: From Words to Meaning**



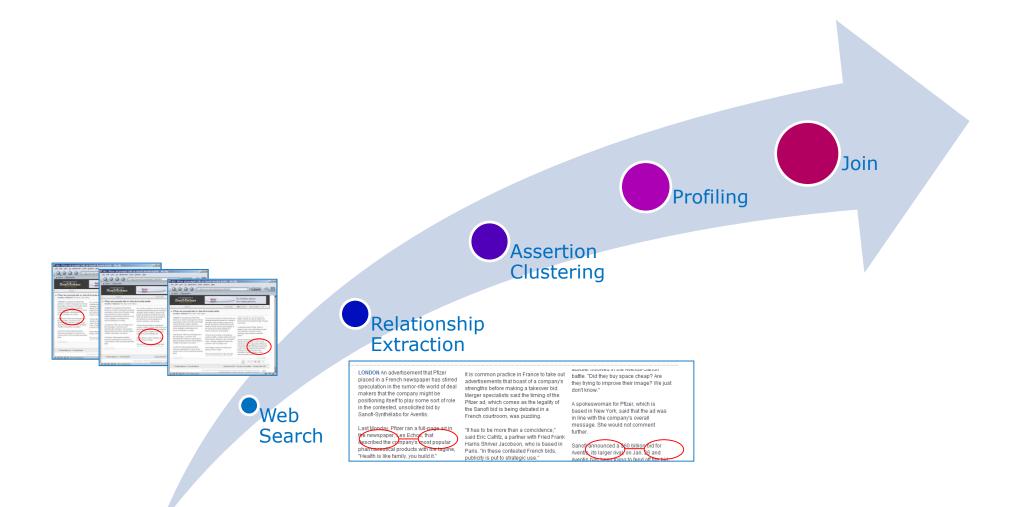


"Among them, nimesulide, a selective COX2 inhibitor, ..."

Consider linguistics to establish relationships

inhibits







# **Linguistics and Ontologies**

- Consider a search for biomarkers for a disease
  - Gene > biomarker > disease

Class1	Relation	Class2		Doc		Hit
▶ ERBB2	prognostic	Breast Neoplasms	▶ 9	19723150	1	Coordinated expression of ER, PR and HER2 define different prognostic subtypes among poorly differentiated breast carcinomas.
▶EGFR	marker	Breast Neoplasms	▶ 3	21276138	1	Currently, CD133 is one of the best markers to characterize cancer stem cells and Her-1 is reported as an important marker for the prognosis of triple-negative breast cancer.
▶ ESR1	predictive	Breast Neoplasms	▶ 4	20437866	1	Molecular subtype approximated by quantitative estrogen receptor, progesterone receptor and Her2 can predict the prognosis of breast cancer.
▶ CD24	marker	Breast Neoplasms	▶ 4	22388757	1	It has been reported previously that: (1) normal-breast epithelial cells that are CD24-/44+ express higher levels of stem/progenitor cell-associated genes; (2) cancer cells that have undergone epithelial to mesenchymal transition display CD24-/44+ cell-surface expression, a marker for breast cancer stem cells; (3) loss of E-cadherin is a preliminary step in epithelial to mesenchymal transition; and (4) vimentin is a marker of mesenchymal phenotype.

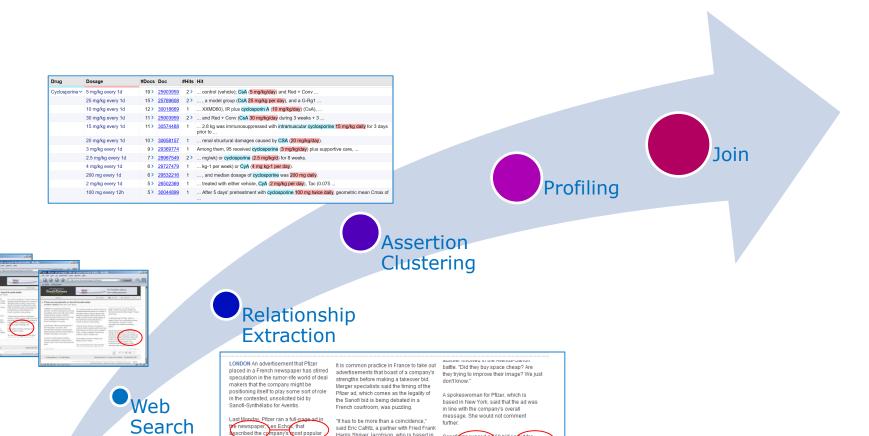


### **Drug Adverse Events**

- What drugs containing a benzene ring and a carboxamide group have a specific adverse event? ChemAxon
  - Substructure search with Jchem in I2E
- Found 3x the number of drugs vs manual curation

Chemicals (Name to S	Stevens-Johnson Synd	FDA Products	Doc	Hit
silantin	Stevens-Johnson Syndrome	PHENYTOIN SODIUM	▶37 7317e374- 847a-4b8a- 851d- 467a0bb9c77e	4Phenytoin Sodium PHENYTOIN dermatitis, lupus erythematosus, Stevens-Johnson syndrome, and toxic epidermal necrolysis
enalaprilat	Stevens-Johnson Syndrome	ENALAPRIL MALEATE	№ 66 759e0a98- 9e32-4caf- 936c- 873ddf8b85b7	8ENALAPRIL MALEATE ENALAPRILAT ANHYDROUS Skin: Exfoliative dermatitis, toxic epidermal necrolysis, Stevens- Johnson syndrome, pemphigus
minocycline	Stevens-Johnson Syndrome	MINOCYCLINE HYDROCHLORIDE	▶68 <u>13d8dea9-5f80-45e6-8a0e-</u> 9b4fd34a560f	6MINOCYCLINE HYDROCHLORIDE MINOCYCLINE of nails, pruritus, toxic epidermal necrolysis, and vasculitis.
glimepiride	Stevens-Johnson Syndrome	GLIMEPIRIDE	▶68 4529ac93- bceb-27fe- e054- 00144ff88e88	6Glimepiride GLIMEPIRIDE anaphylaxis, angioedema, and Stevens-Johnson Syndrome [see





Sanof announced a \$60 billion old for Aventic its larger rival on Jan. 86 and

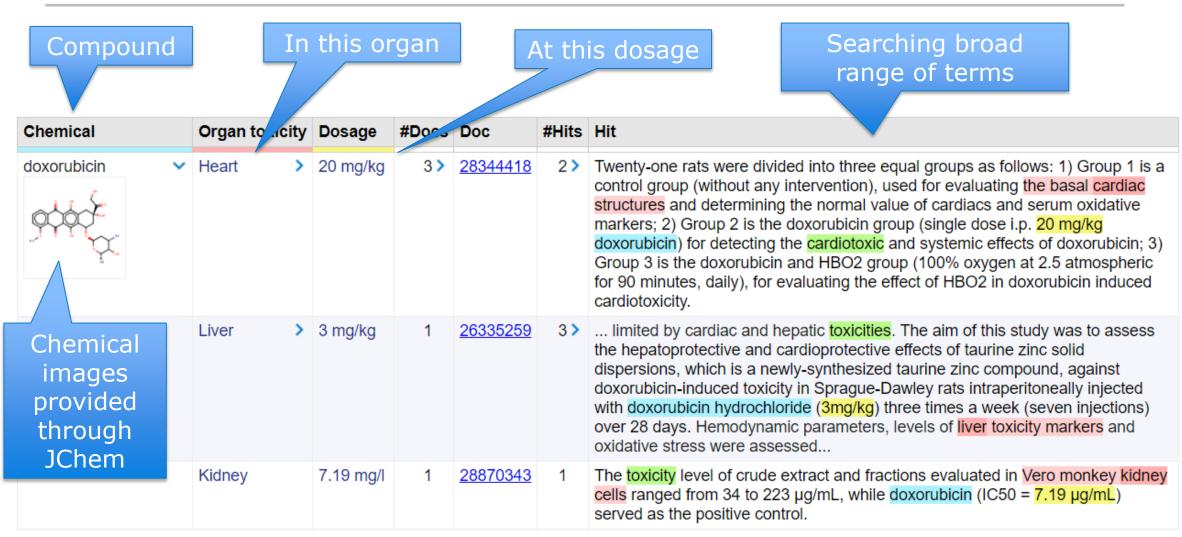
Harris Shriver Jacobson, who is based in

Paris. "In these contested French bids, publicity is put to strategic use.

pharmaceutical products with the tagline,



### **Assertion Extraction: Potential Safety Issues**













LONDON An advertisement that Pfizer placed in a French newspaper has stirred makers that the company might be positioning itself to play some sort of role Web in the contested, unsolicited bid by Sanofi-Synthélabo for Aventis.

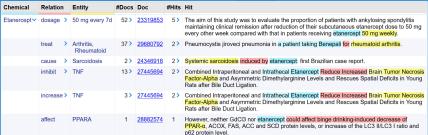
day. Pfizer ran a full-page the newspaper, es Echos, that described the company's most popular pharmaceutical products with the tagline, It is common practice in France to take out battle. "Did they buy space cheap? Are advertisements that boast of a company's they trying to improve their image? We just strengths before making a takeover bid. Merger specialists said the timing of the Pfizer ad, which comes as the legality of

publicity is put to strategic use.

A spokeswoman for Pfizer, which is the Sanofi bid is being debated in a based in New York, said that the ad was French courtroom, was puzzling. in line with the company's overall message. She would not comment "It has to be more than a coincidence," said Eric Cafritz, a partner with Fried Frank

Harris Shriver Jacobson, who is based in Paris. "In these contested French bids, Aventis, its larger rival on Jan. 26 and





Join



Search

# **Chemical Information**

For a given drug					ort var	ious	properties			
	Chenical	Relation	Entity	#Docs	Doc	#Hits	Hit			
	brivaracetam V	dosage >	200 mg every 12h	3>	27842261	1	Carbamazepine epoxide plasma concentrations were approximately doubled by brivaracetam 100 or 200mg bid.			
	ges ided ugh	treat 💙	Seizures	8>	29270790	1	Budget Impact Analysis of Brivaracetam Adjunctive Therapy for Partial- Onset Epileptic Seizures in Valencia Community, Spain.			
			Epilepsy	3 >	<u>26760311</u>	1	Brivaracetam for the treatment of epilepsy.			
Chem			Epilepsies, Partial	2>	28829199	1	The preclinical discovery and development of brivaracetam for the treatment of focal epilepsy.			
imag			Depressive Disorder	1	<u>19211275</u>	1	Brivaracetam inhibits spreading depression in rat neocortical slices in vitro.			
provid throu JChe		affect	SV2A	1	<u>26899665</u>	1	Brivaracetam binds to synaptic vesicle glycoprotein 2a (SV2A) in the brain with greater selectivity and 15- to 30-fold higher affinity than levetiracetam, as demonstrated in preclinical models, and has demonstrated efficacy in reducing the frequency of partial onset seizures in clinical trials.			
		IC50	8.2e-06 mol/l	1	<u>25986188</u>	1	In vitro hydrolysis of carbamazepine-epoxide in human hepatocytes was inhibited by brivaracetam, with an <a href="IC50">IC50</a> of <a href="8.2\mu M">8.2\mu M</a> .			
		inhibit	hydrolase	1	<u>27842261</u>	1	Valproate and brivaracetam inhibit epoxide hydrolase and increase carbamazepine epoxide levels.			

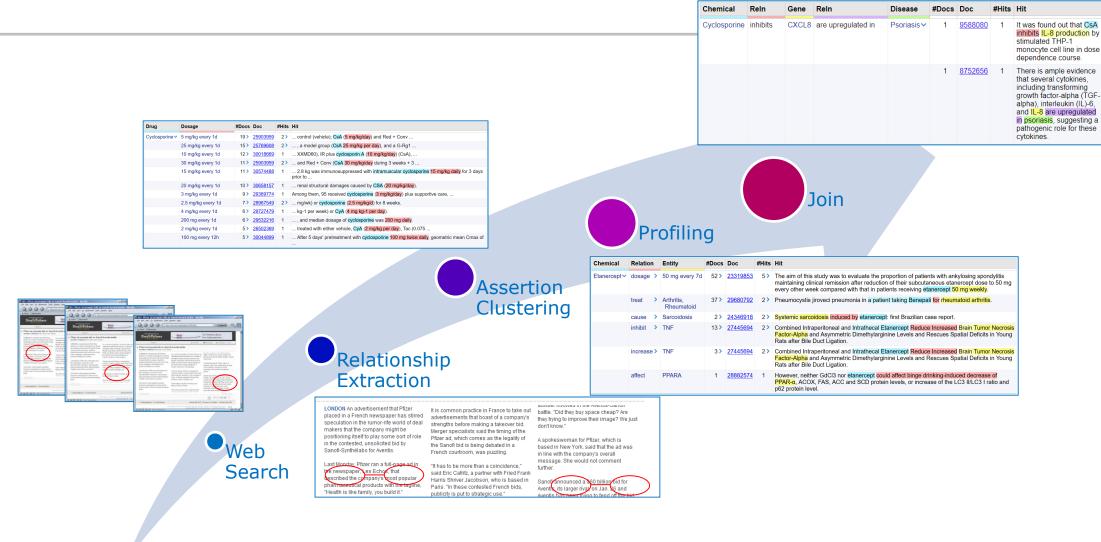


# **Extracting Product Information for IDMP**

#### IDMP (IDentification of Medicinal Products) ISO standard

Source	Section	Full Name	Invented Name Part	Strength Part	Pharmaceutical Dose	Intended Use Part	Flavour Part	#Hits	Hit
SmPC	1	Abacavir/Lamivudine 600 mg/300 mg film-coated tablets	Abacavir/Lamivudine	600 mg/300 mg	film-coated tablets			3	Abacavir/Lamivudine 600 mg/300 mg film- coated tablets
SmPC	1	ABILIFY 7.5 mg/ml solution for injection	ABILIFY	7.5 mg/ml	solution for injection			3	ABILIFY 7.5 mg/ml solution for injection
SmPC	1	Accupro Tablets 5 mg	Accupro	5 mg	Tablets			3	Accupro Tablets 5 mg
SmPC	1 💙	Acidex Liquid - peppermint flavour	Acidex		Liquid		peppermint flavour	3	Acidex Liquid - peppermint flavour
		Boots Heartburn Relief Peppermint Flavour	Boots			Heartburn Relief	Peppermint Flavour	3	Boots Heartburn Relief Peppermint Flavour
		Healthpoint Heartburn and Indigestion Oral Suspension	Healthpoint		Oral Suspension	Heartburn and Indigestion		3	Healthpoint Heartburn and Indigestion Oral Suspension
		LloydsPharmacy Heartburn and Indigestion Liquid Peppermint Flavoured	LloydsPharmacy		Liquid	Heartburn and Indigestion	Peppermint Flavoured	3	LloydsPharmacy Heartburn and Indigestion Liquid Peppermint Flavoured
		Peptac Peppermint Liquid	Peptac		Liquid		Peppermint	3	Peptac Peppermint Liquid
		Tesco Heartburn & Indigestion relief	Tesco			Heartburn & Indigestion relief		3	Tesco Heartburn & Indigestion relief



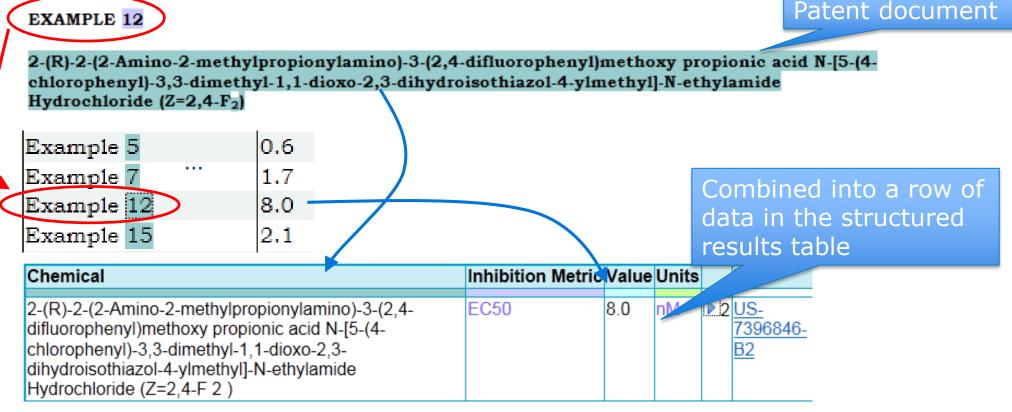




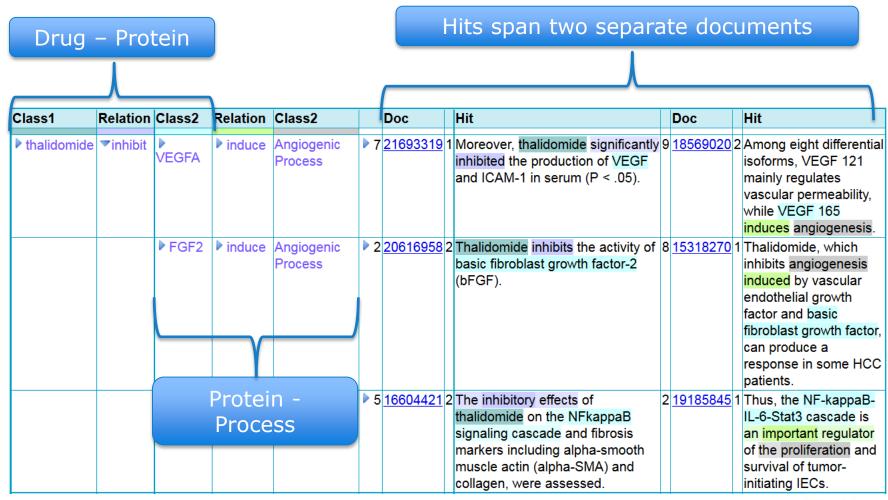
### **Linking Information within Documents**

Connecting information found in different parts of the document

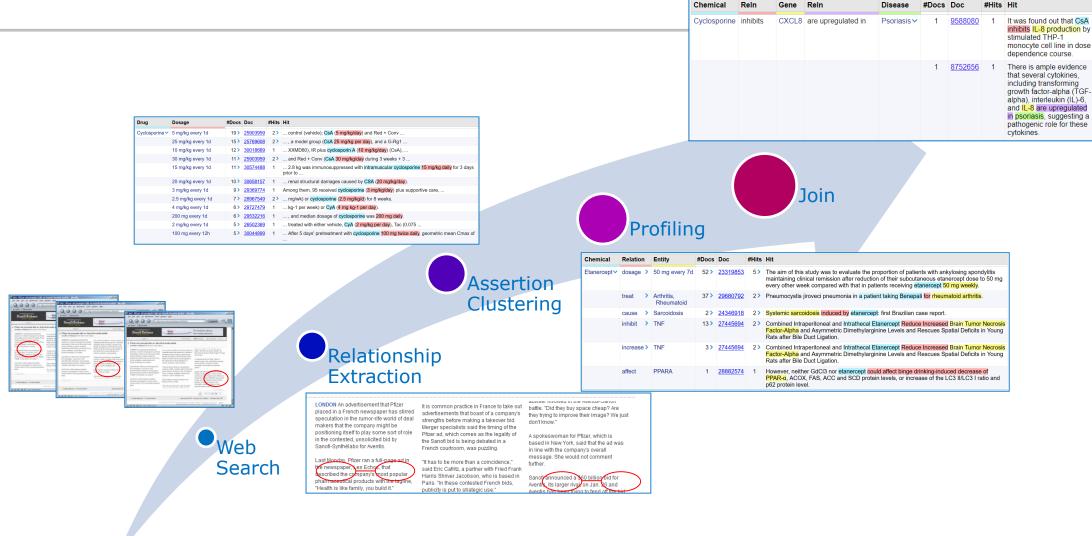
For example finding a compound as "Example 12" in a patent and linking to a table where numerical data is reported



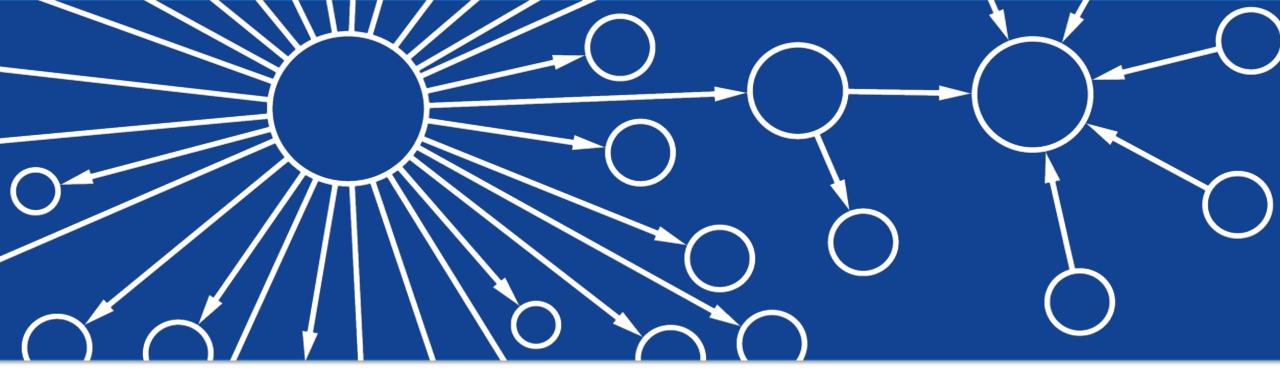
#### **Connecting Information Between Documents**











# Thank you!

For more information...

Visit: www.linguamatics.com

Contact: jeff.nauss@linguamatics.com

Healthcare Blog

Life Sciences Blog







