



## Integrating ChemAxon and Linguamatics to provide Agile, Chemistry-enabled Text Mining

Jeffrey L. Nauss, PhD  
Application Specialist, Linguamatics

ChemAxon UGM, September 2010

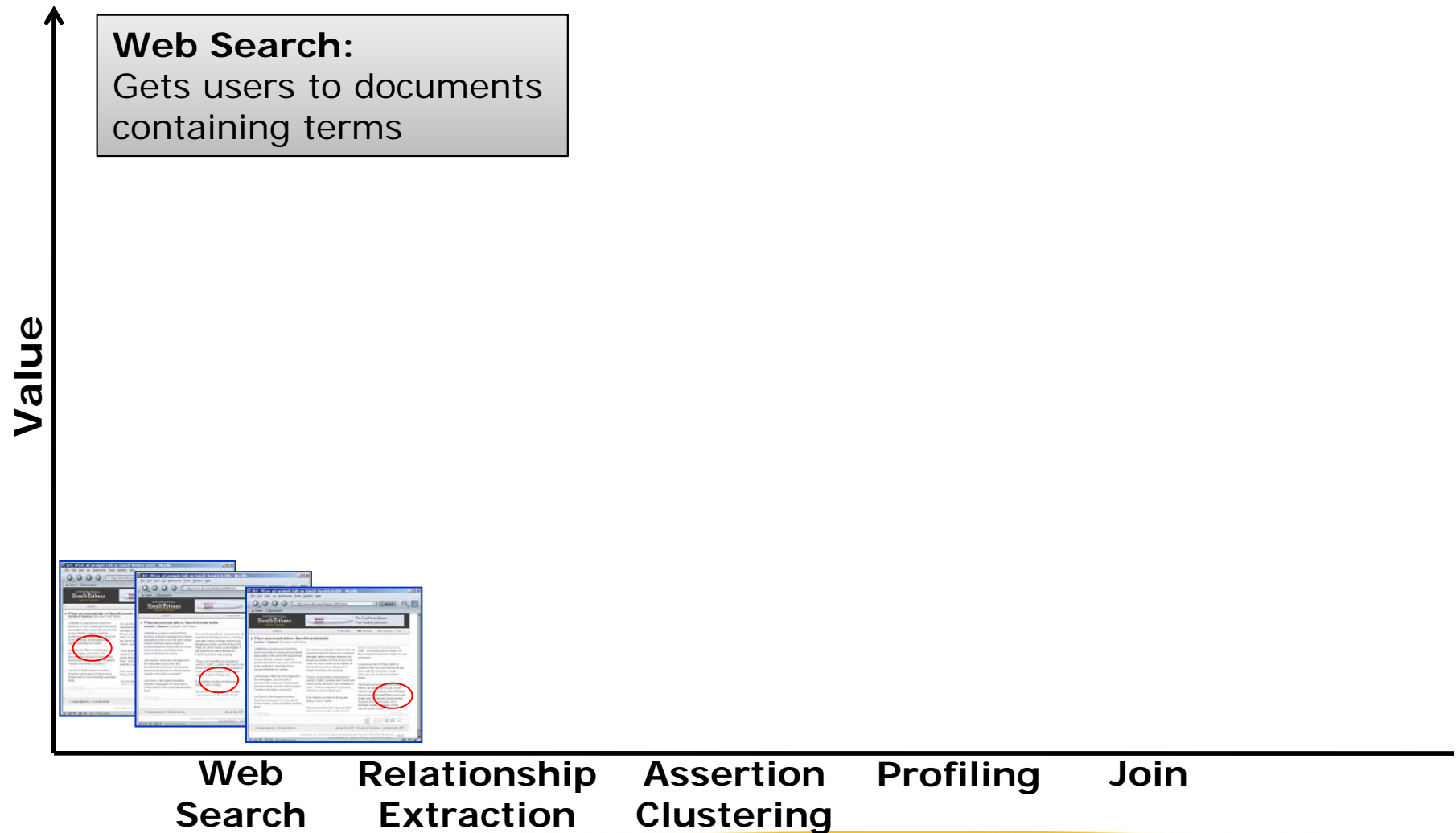


# Linguamatics I2E

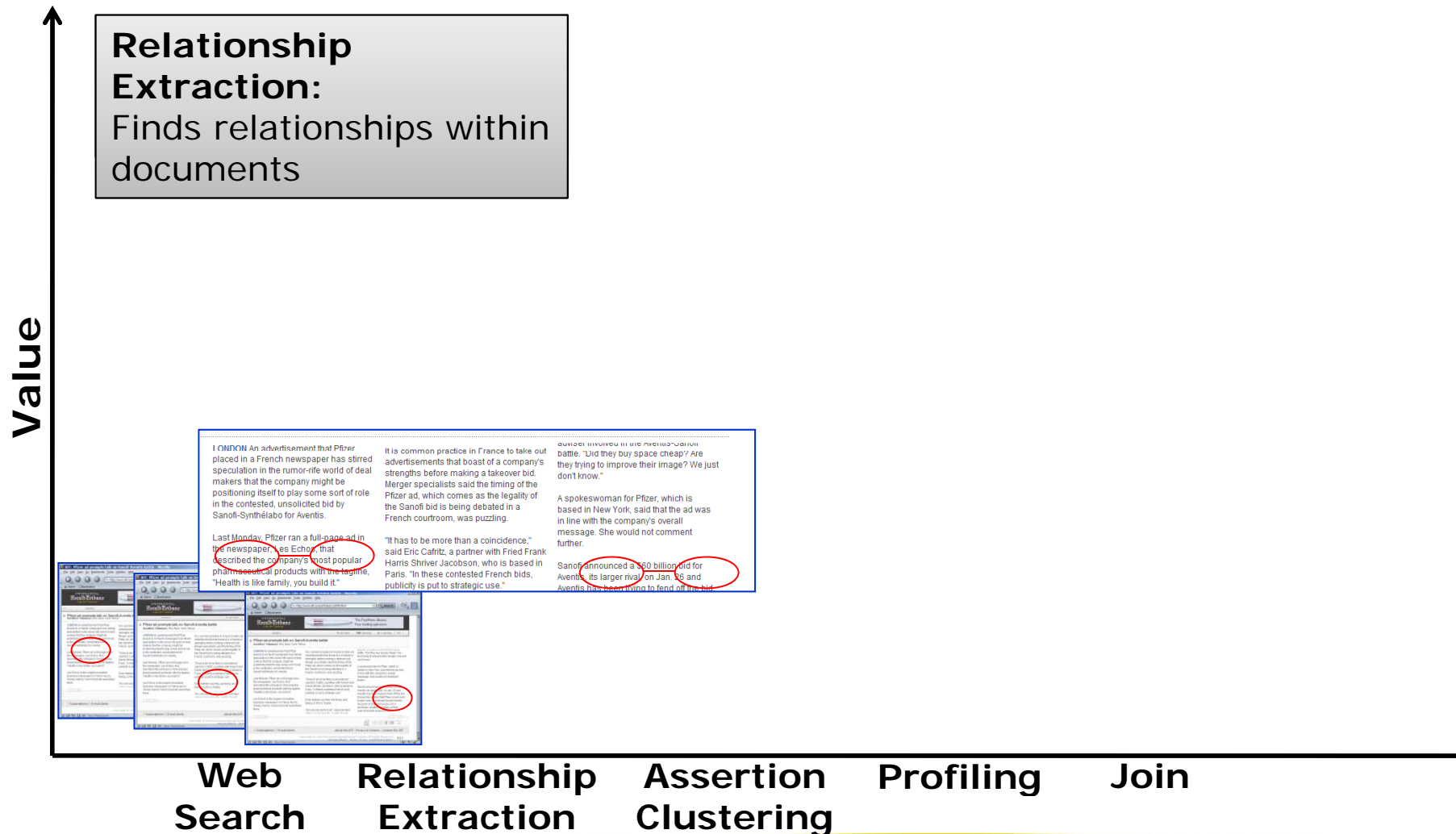
## **Flexible and scalable text mining for business-critical knowledge discovery**

- ◆ Natural language processing (NLP)-based knowledge discovery platform
- ◆ Rapidly reveals structured facts and relationships by understanding meaning
- ◆ Delivers relevant, high quality results in real-time

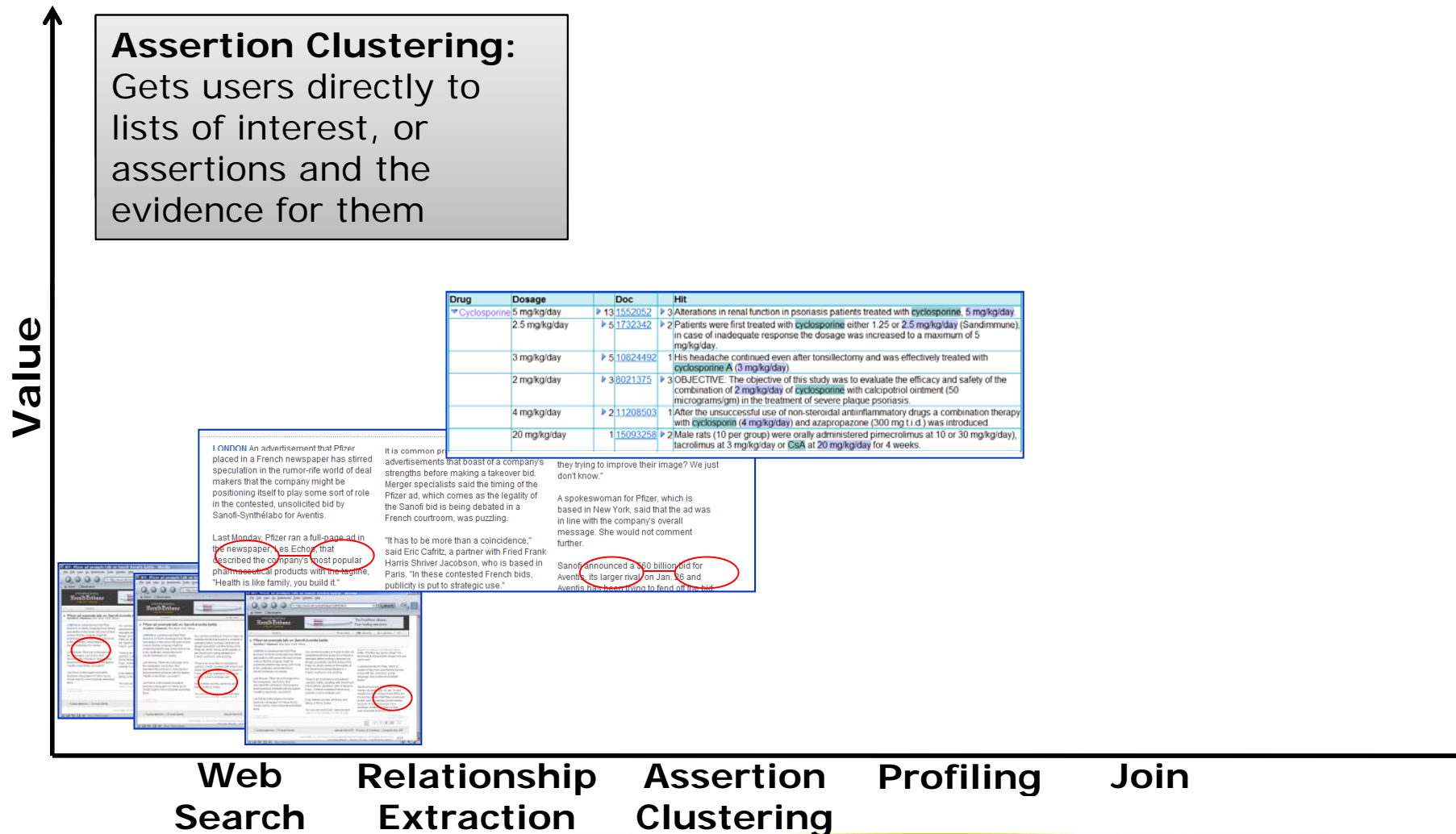
# From Documents to Knowledge



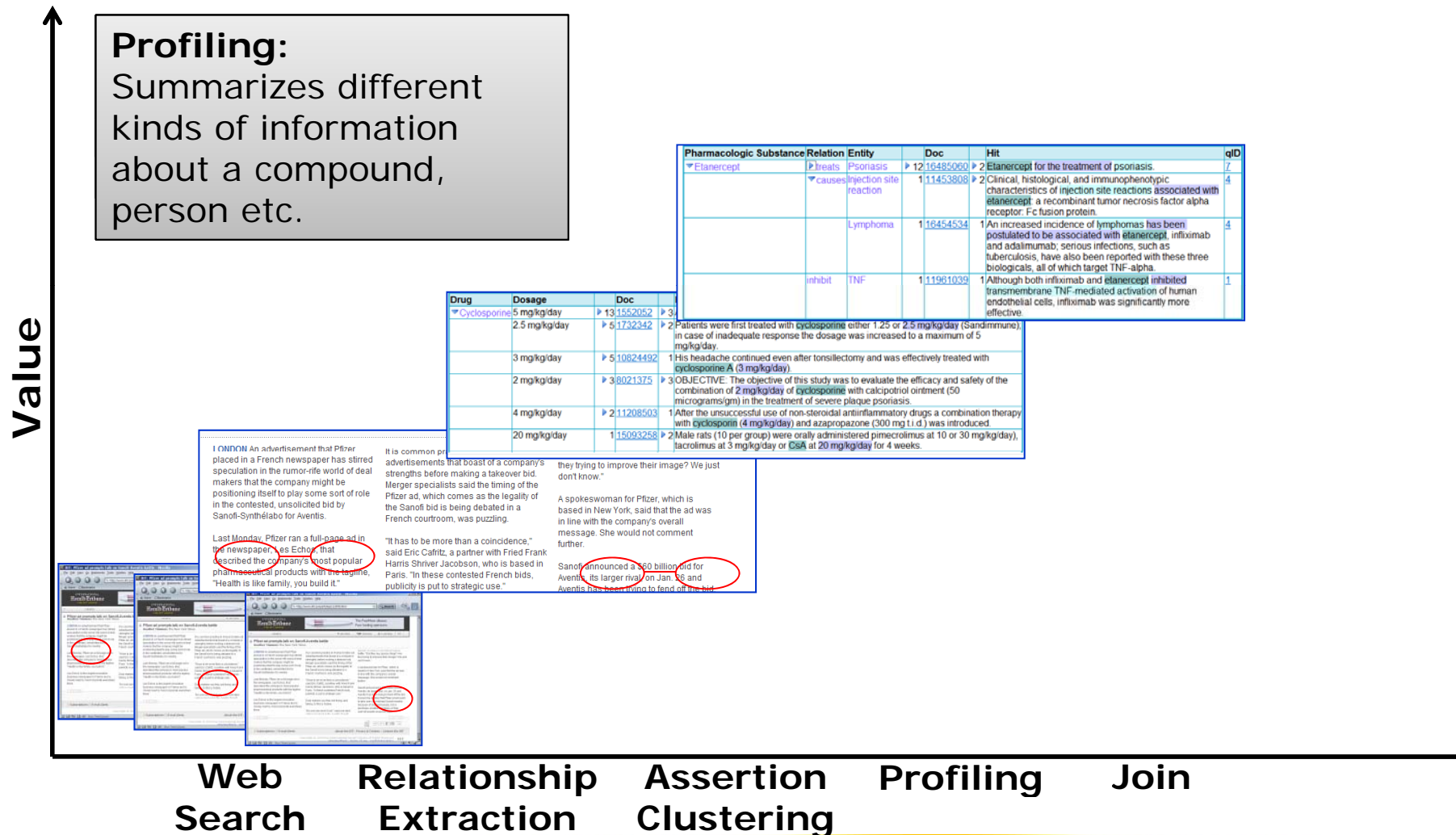
# From Documents to Knowledge



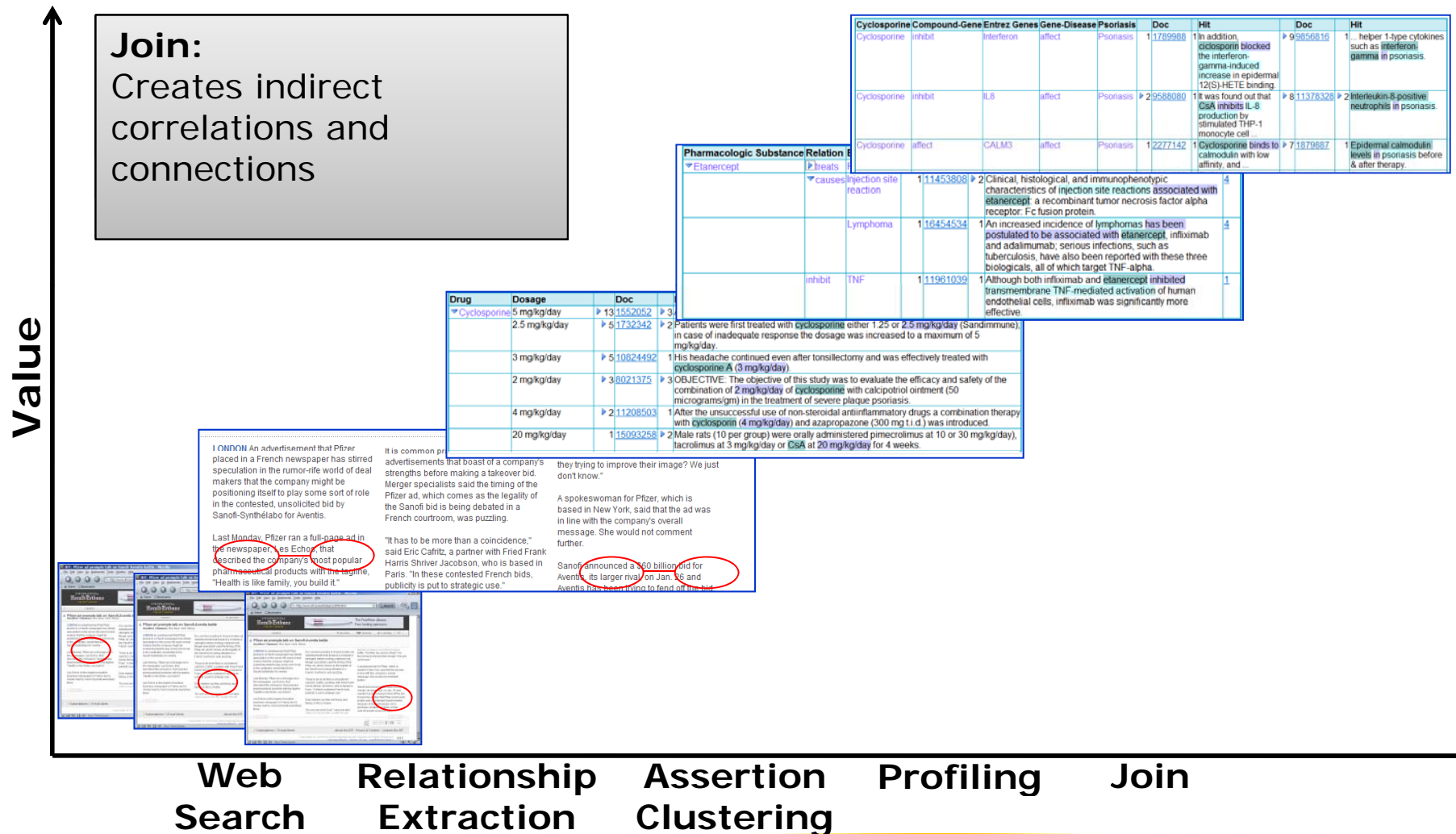
# From Documents to Knowledge



# From Documents to Knowledge



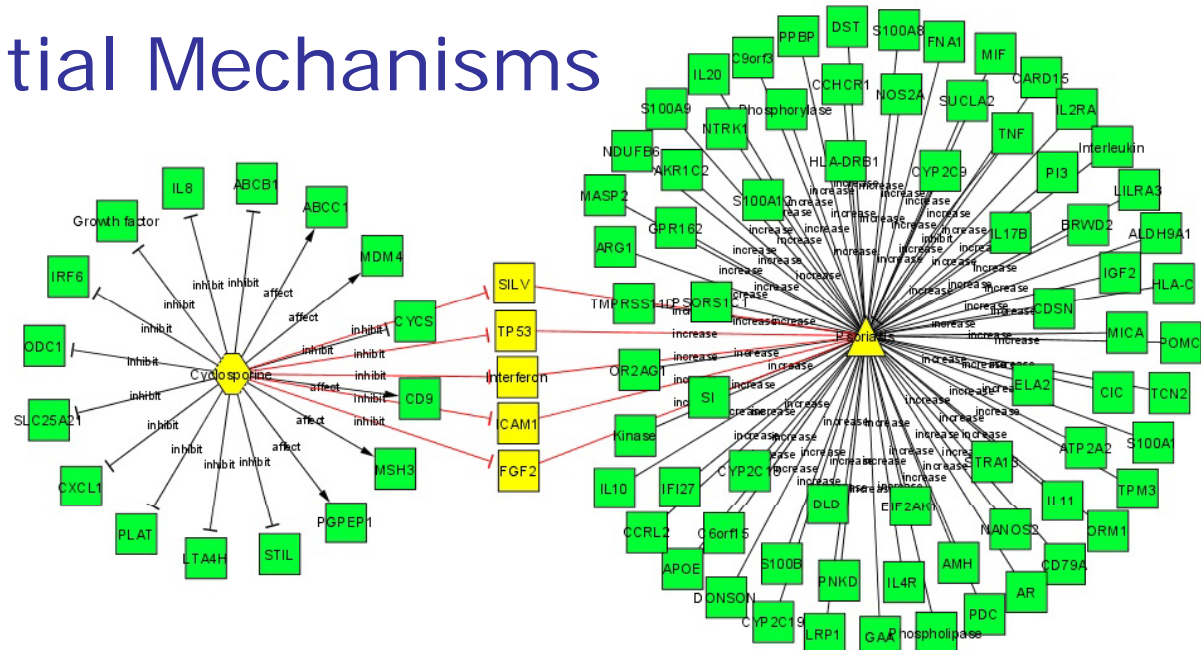
# From Documents to Knowledge



Web Search      Relationship Extraction      Assertion Clustering      Profiling      Join

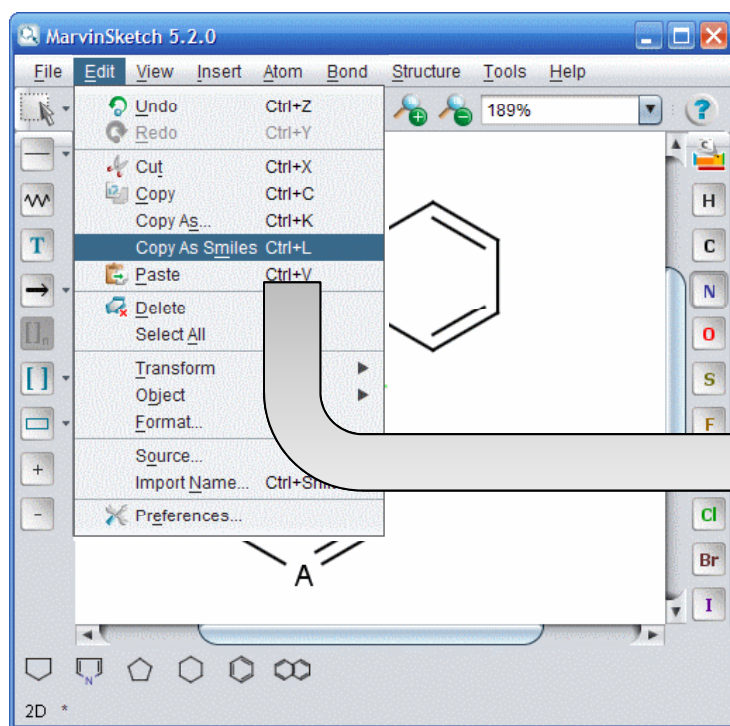
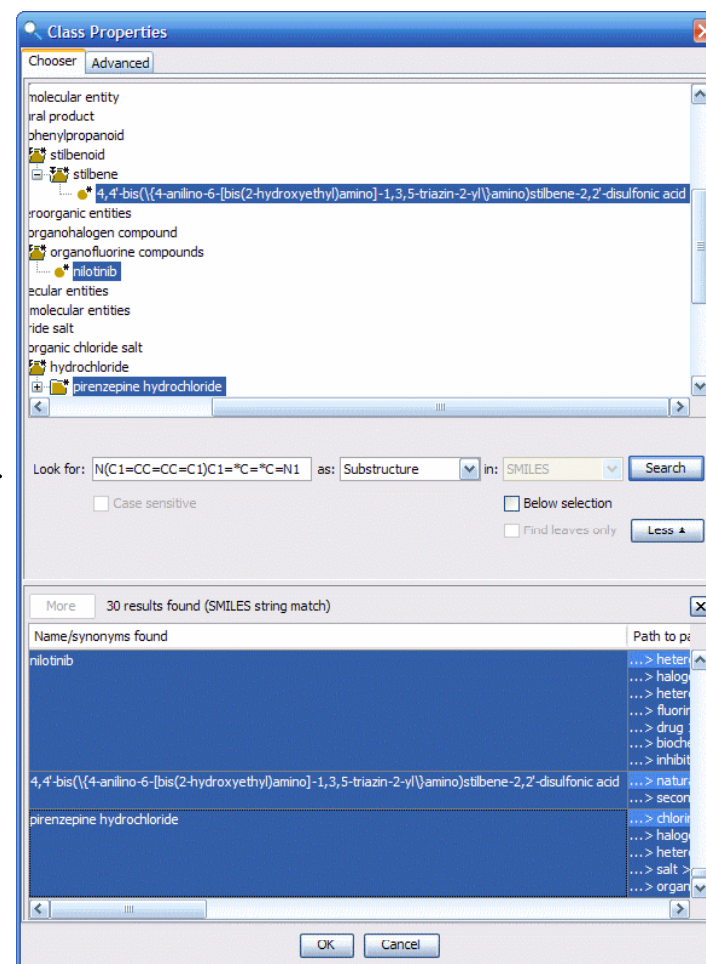
# Finding Potential Mechanisms

Discover indirect associations across multiple documents



Cyclosporine	Compound-Gene	Entrez Genes	Gene-Disease	Psoriasis	Doc	Hit	Doc	Hit
Cyclosporine	inhibit	Interferon	affect	Psoriasis	1 <a href="#">1789988</a>	1 In addition, ciclosporin blocked the interferon-gamma-induced increase in epidermal 12(S)-HETE binding.	▶ 9 <a href="#">9856816</a>	1 ... helper 1-type cytokines such as interferon-gamma in psoriasis.
Cyclosporine	inhibit	IL8	affect	Psoriasis	▶ 2 <a href="#">9588080</a>	1 It was found out that CsA inhibits IL-8 production by stimulated THP-1 monocyte cell ...	▶ 8 <a href="#">11378328</a>	▶ 2 Interleukin-8-positive neutrophils in psoriasis.
Cyclosporine	affect	CALM3	affect	Psoriasis	1 <a href="#">2277142</a>	1 Cyclosporine binds to calmodulin with low affinity, and ...	▶ 7 <a href="#">1879887</a>	1 Epidermal calmodulin levels in psoriasis before & after therapy.
Cyclosporine	inhibit	Growth factor	affect	Psoriasis	▶ 2 <a href="#">8884530</a>	1 FK506 and cyclosporin A inhibit growth factor-stimulated human keratinocyte proliferation by blocking cells in the ...	▶ 6 <a href="#">14962110</a>	1 Single-nucleotide polymorphisms of vascular endothelial growth factor in psoriasis of early onset.

# Extract Information Based on Chemical Structure

Class Properties

Chooser Advanced

molecular entity  
oral product  
phenylpropanoid  
stilbenoid  
stilbene  
4,4-bis((4-anilino-6-[bis(2-hydroxyethyl)amino]-1,3,5-triazin-2-yl)amino)stilbene-2,2'-disulfonic acid  
inorganic entities  
organohalogen compound  
organofluorine compounds  
nilotinib  
molecular entities  
molecular entities  
oxide salt  
organic chloride salt  
hydrochloride  
pirenzepine hydrochloride

Look for: N(C1=CC=CC=C1)C1=CC=C(N1) as: Substructure in: SMILES Search

Case sensitive  Below selection  Find leaves only Less A

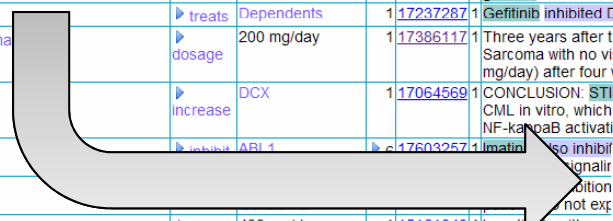
More 30 results found (SMILES string match)

Name/synonyms found	Path to p...
nilotinib	...> hetero ...> halogen ...> hetero ...> fluorine ...> drug ...> biochem ...> inhibit
4,4-bis((4-anilino-6-[bis(2-hydroxyethyl)amino]-1,3,5-triazin-2-yl)amino)stilbene-2,2'-disulfonic acid	...> natural ...> second
pirenzepine hydrochloride	...> chlorine ...> halogen ...> hetero ...> salt ...> organ

OK Cancel

# Link Results to Outside Sources

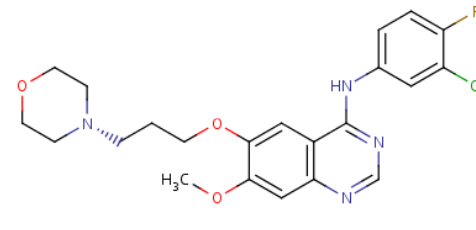
Pharmacologic Substance	Targets	Entrez Genes	Doc	Hit
▼ gefitinib	▶ inhibit	AHSA1	1 17205515	Inversely, in IC1LC131, Erk and Akt pathways remained active, while Jnk and P38 pathways were inhibited by gefitinib.
	▶ treats	Dependents	1 17237287	Gefitinib inhibited DNA synthesis in a concentration-dependent fashion in 6 of 17 lines.
▼ imatinib	▶ dosage	200 mg/day	1 17386117	Three years after treatment with different chemotherapeutic agents for progressive cutaneous Kaposi's Sarcoma with no visceral involvement, he was prescribed imatinib (200 mg/day for two weeks followed by 400 mg/day) after four weeks of treatment he developed anasarca, further progression of KS and agranulocytosis.
	▶ increase	DCX	1 17064569	CONCLUSION: STI571 could promote the activation/maturation of DC derived from BMMNCs of patients with CML in vitro, which might be partially responsible for the fact that the inhibitory effect of VEGF on DC NF- $\kappa$ B activation was relieved through STI571 inhibiting the overproduction of VEGF in CML.
	▶ inhibit	ABL1	1 17603257	Imatinib also inhibited the phosphorylation of ABL1 in the presence of imatinib.
▼ imatinib methanesulfonate	▶ dosage	400 mg/day	1 15161340	In patients with new the chronic phase, resulted in higher progression to the
	▶ increase	AIF1	1 16885745	Gleevec alone resist G2 arrest.
	▶ inhibit	ABL1	▶ 2 17436575	Gleevec inhibits the
	▶ treats	Adhesion	1 17289809	Depletion of endoc of EPCs to activate
▼ lapatinib	▶ inhibit	AKT1	1 17283152	In addition, lapatinib non-overexpressin
	treats	Insulin-like growth factor	1 17308062	Importantly, lapatinib parental and resist blocking antibody
▼ nilotinib	▶ inhibit	ABL1	1 17106016	Nilotinib inhibits BC
	treats	International normalised ratio	1 17106016	Nilotinib inhibits BC
▼ vandetanib	▶ increase	ABCB1	1 17912240	In addition, ZD6474 by enhancing the t
	▶ inhibit	EGFR	▶ 2 17308046	In vitro, ZD6474 inhib and VEGF-induced



chemicalize.org<sup>alpha</sup> Open All Close All Manage calculations Layout: Medicinal Chemist

gefitinib Search

**Molecule**



**logP** ? -

logP: 3.88

**Polar Surface Area** ? -

Polar surface area: 68.74

**Molecular Surface Area** ? -

Calculate Molecular Surface Area

**Name** ? -

IUPAC name: N-(3-chloro-4-fluorophenyl)-7-methoxy-6-[3-(morpholin-4-yl)propoxy]quinazolin-4-amine

Traditional name: Iressa

**Elemental Analysis** ? -

Formula: C<sub>22</sub>H<sub>24</sub>ClFN<sub>4</sub>O<sub>3</sub>  
 Isotope formula: C<sub>22</sub>H<sub>24</sub>ClFN<sub>4</sub>O<sub>3</sub>  
 Composition: C (59.13%), H (5.41%), Cl (7.93%), F (4.25%), N (12.54%), O (10.74%)  
 Isotope composition: C (59.13%), H (5.41%), Cl (7.93%), F (4.25%), N (12.54%), O (10.74%)  
 Mass: 446.902  
 Exact mass: 446.152096566

**Topology Analysis** ? -

Simple	Ring Counts	Path and distance
Atom count: 55	Bond count: 58	Cyclomatic number: 4
Chain atom count: 9	Chain bond count: 11	Asymmetric atom count: 0
	Rotatable bond count: 8	

# Thank You!

## For more information...

**Visit:** [www.linguamatics.com](http://www.linguamatics.com)

**Contact:** Susan LeBeau

Email: [susan.lebeau@linguamatics.com](mailto:susan.lebeau@linguamatics.com)

Phone: 774-571-1117

Mobile: 774-277-1256

## Webinars:

Visit <http://www.linguamatics.com/welcome/events/webinars.html>

## Meet our experts at upcoming events in 2010:

Visit <http://www.linguamatics.com/welcome/events/conferences.html>