

Linguamatics Update

ICIC 2012

Berlin

October 2012

About Linguamatics



I2E: agile, scalable, real-time NLP-based text mining
Fact extraction and knowledge synthesis

Fortune 500
Pharma/Biotech

Healthcare
Government

Including 9
of the top 10

Whatever the Content...

- ... I2E can mine and extract with precision

Scientific literature

Gene/Protein	Biomarker	Disease	Doc	Hit
▼ TRIM25	as a new potential biomarker for	Breast cancer	1 16144914	1 Estrogen-responsive finger protein as a new potential biomarker for breast cancer.
	is a significant prognostic factor in	Breast cancer	1 16144914	1 CONCLUSIONS: Our data suggest that Efp immunoreactivity is a significant prognostic factor in breast cancer patients.
	was significantly correlated with	Breast cancer	1 16144914	1 Moreover, Efp immunoreactivity was significantly correlated with poor prognosis of breast cancer patients, and multivariate analyses of disease-free survival and overall survival for 151 breast cancer patients showed that Efp immunoreactivity was the independent marker.
▼ BCL2	is a prognostic marker in	Breast cancer	1 16638854	1 Bcl-2 is a prognostic marker in breast cancer independently of the Nottingham Prognostic Index.
	is an independent predictor of	Breast cancer	1 16638854	1 CONCLUSION: Bcl-2 is an independent predictor of breast cancer outcome and seems to be useful as a prognostic adjunct to the NPI, particularly in the first 5 years after diagnosis.

Twitter

Intention	Treatment	User	Doc	Hit
▶ Just got	▶ my flu shot	@4momama	1 217	1 @4momama... Just got my flu shot and in meeting for the ...
▶ I got	▶ my flu shot	@lilii89	▶ 2 657	1 @lilii89... I got my flu shot today!
▶ got	▶ my flu shot	@aaaaaaannie	1 658	1 @aaaaaaannie... got my flu shot but now dizzy from the ...
▼ Getting	▶ my flu shot	@amytnguyen	1 737	1 @amytnguyen... Getting my flu shot on
	▶ a flu shot	@oregonian	▶ 3 650	1 @oregonian... Getting a flu shot (or two)?
	▶ flu shot	@bobbartz	1 767	1 @bobbartz... Getting flu shot
	▶ the flu shot	@akronbabygirl	1 262	1 @akronbabygirl... Getting the flu shot?
	▶ flu shots	@lauraann9	1 180	1 @lauraann9... Getting flu shots and having blood drawn is ...
	▶ my free flu shot	@brittag10	1 752	1 @brittag10... Getting my free flu shot!

Recent Customer News

- Continually growing customer base...

LINGUAMATICS' I2E TEXT MINING PLATFORM CHOSEN BY FDA

CDER to use Linguamatics I2E in support of drug safety research activities

(Cambridge, England and Boston, USA – August 1st 2012)

Linguamatics, the leader in natural language processing (NLP)-based text mining, announced today that the FDA's Center for Drug Evaluation and Research (CDER) has licensed its I2E text mining platform to support laboratory research efforts on drug safety...

See: www.linguamatics.com/welcome/news/press_releases.html

Just Launched: I2E Patent Solution

I2E text mining

- High performance NLP querying
- Fact extraction, knowledge synthesis
- Use of domain thesauri
- Ability to cross document silos

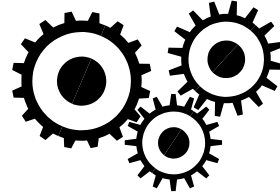
Hosted platform

- Reduced Cost of Ownership
- No installation delay

Full text patents

- 17 million documents (USPTO, EPO, WIPO)
- Applications and Grants
- From IFI Claims Direct

Coming Soon: New Web Services API



I2E Server

Supports improved automation and integration, at scale

Standard GUI

Web services API

The screenshot shows the I2E Express Standard GUI. At the top, there's a search bar with the text "MEDLINE Sep 2012 with up..." and "for biomarker". Below the search bar, there are tabs for "Standard Query", "Smart Query", and "Pro". On the left side, there are links for "Help", "Express Tips", "FAQ", "Local Help", and "Local Support". The main content area is titled "Biomarker by disease" and contains a form with "Gene" and "Disease" dropdown menus. Below the form, there's a paragraph of text: "This query finds biomarker relations using the Biomarker macro. An example relation is 'prognostic in'." and "An example hit might be: Her-2/neu overexpression is an important prognostic parameter in breast cancer patients." At the bottom, there are buttons for "+Class", "+Alt List", and "Restore defaults".

The screenshot shows the I2E Express Web Services API interface. It features a "Linguamatics" logo at the top left and tabs for "Chemicals", "Diseases", and "Genes/Proteins". The main content area is titled "Most Frequent Chemicals" and displays a pie chart with segments for "Gefitinib", "Vorinostat", "Recombinant Vascular Endothelial Growth Factor", "MEK Inhibitor AZD6244", "Adjuvant", and "Pharmacologic Substance". Below the pie chart, there's a text box with a document icon and the following text: "ERCC2.Lung neoplasm malignant Doc #18 The presence of the variant alleles 312Asn and 751Gln of XPD have been associated with relatively high risk of lung cancer in Caucasian [17-20] and Asian [21-24] populations and a recent meta-analysis concludes that the variant genotypes 312Asn/Asn and 751Gln/Gln are associated with a statistically significant lung cancer risk in the Caucasian population [25]. ...Additionally, interaction between XPC and XPD polymorphisms showed an increased risk of lung cancer (OR = 2.25). ... Doc #86 The Asp/Asp genotype of the ERCC2/XPB Asp312Asn polymorphism was found to have an increased risk of lung cancer when the combined Asp/Asn and Asn/Asn genotypes served as reference (OR = 1.86, 95% CI = 1.02 - 3.40) in Polish men [20]. ...Table 4 shows the association between the ERCC2.Lung312Asn polymorphism and lung cancer." To the right of the text box is a network graph with nodes and edges, representing relationships between various entities like "Neoplasm malignant", "Recombinant Vascular Endothelial Growth Factor", "Gefitinib", "Vorinostat", etc.

Coming Soon: Chemistry with Name to Structure

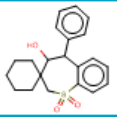
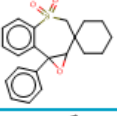
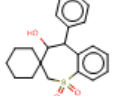
- Integrates ChemAxon Name to Structure

- Substructure and Similarity Search

- Extraction of properties

- Output to e.g. Microsoft Excel

Chemical	Feature	Value	Doc	Hit
4-hydroxy-5-phenyl-4,5-dihydro-2H-spiro[1λ ⁶ -benzothiepine-3,1'-cyclohexane]-1,1-dione	mp	209-210°	US-6943189-B2	1 Example 31 cis-4-Hydroxy-5-phenyl-2,3,4,5-tetrahydro spiro(benzothiepine-3,1'-cyclohexane)-1,1-dioxide (60) ... mg of white crystal, mp 209-210° C. Proton and carbon NMR ...
2'-phenyl-3'-oxa-7λ ⁶ -thiaspiro [cyclohexane-1,5'-tricyclo [6.4.0.0 ^{2,4}] dodecane]-1'(8'),9',11'-triene-7',7'-dione	mp	154-155°	US-6642268-B2	1 8b-Phenyl-1a,2,3,8b-tetrahydrospiro (benzothiepine[4,5-b]oxirene-2,1'-cyclohexane)-4,4-dioxide (58) ...) of yellow solid, mp 154-155° C. Proton and carbon NMR ...
4-hydroxy-5-phenyl-4,5-dihydro-2H-spiro[1λ ⁶ -benzothiepine-3,1'-cyclohexane]-1,1-dione	mp	99-100°	US-6642268-B2	1 trans-4-Hydroxy-5-phenyl-2,3,4,5-tetrahydro spiro(benzothiepine-3,1'-cyclohexane)-1,1-dioxide (59) ... as a white solid, mp 99-100° C. Proton NMR showed this ...

Chemical	Feature	Value	Hit
	mp	209-210°	Example 31 cis-4-Hydroxy-5-phenyl-2,3,4,5-tetrahydro spiro(benzothiepine-3,1'-cyclohexane)-1,1-dioxide (60) ... mg of white crystal, mp 209-210° C. Proton and carbon NMR ...
	mp	154-155°	8b-Phenyl-1a,2,3,8b-tetrahydrospiro(benzothiepine[4,5-b]oxirene-2,1'-cyclohexane)-4,4-dioxide (58) ...) of yellow solid, mp 154-155° C. Proton and carbon NMR ...
	mp	99-100°	trans-4-Hydroxy-5-phenyl-2,3,4,5-tetrahydro spiro(benzothiepine-3,1'-cyclohexane)-1,1-dioxide (59) ... as a white solid, mp 99-100° C. Proton NMR showed this ...

Thank You

For more information...

Visit our stand -

To see live demos and talk to our experts

Talk tomorrow -

Text Mining Patents at Scale: Challenges and Successes

Visit: www.linguamatics.com

Contact:

Email: enquiries@linguamatics.com

Tel: +44 (0)1223 651910