

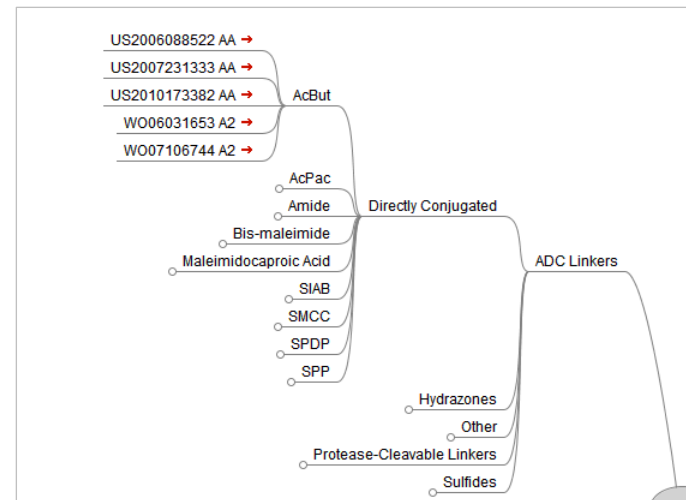
Exploring Patent Landscapes with Linguamatics I2E

Piotr Masiakowski

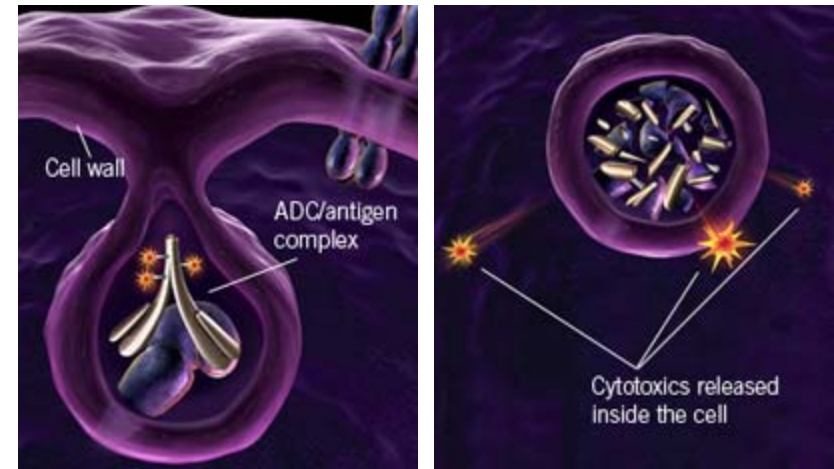
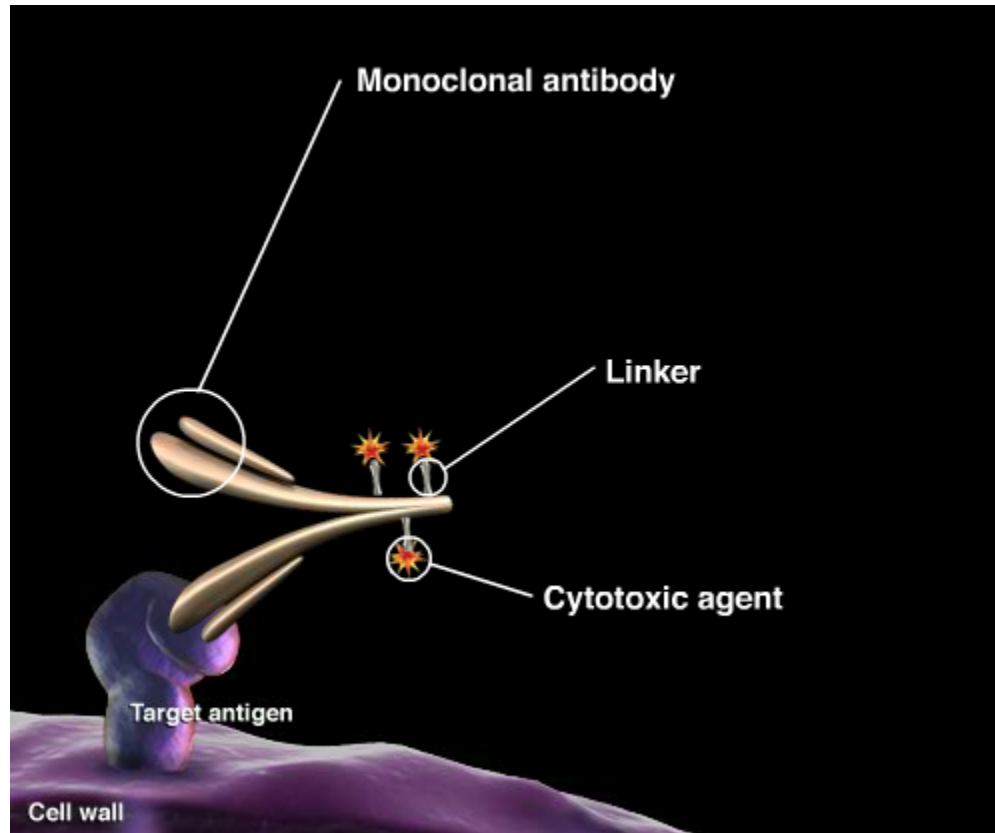
Linguamatics Fall Users Conference 2011

AGENDA

- What is a patent landscape? – Examples
- Why is I2E a great landscaping tool?
- Problems with text-mining patents
- Our approach to reviewing and editing I2E results
- Putting it all together



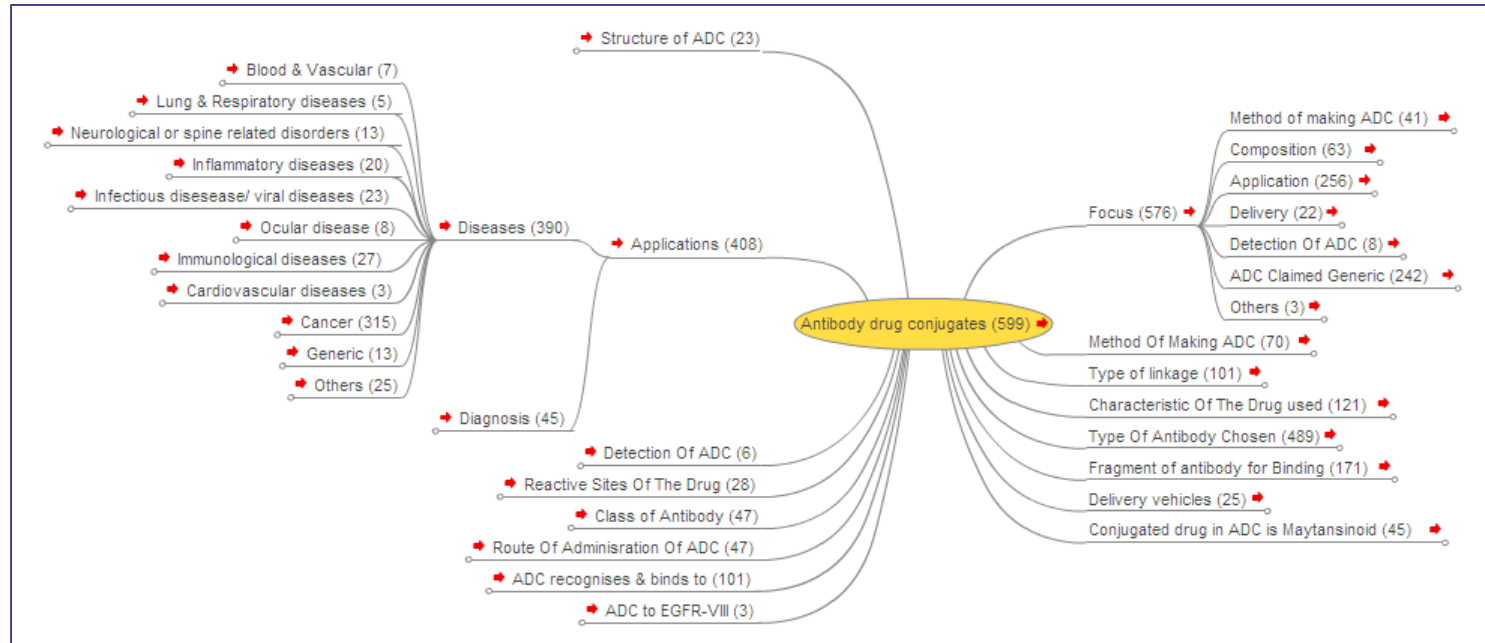
ANTIBODY-DRUG CONJUGATES OVERVIEW



- A promising technology for delivering powerful toxins selectively to cancer cells
- Thousands of patents and patent applications focusing on various aspects of this technology

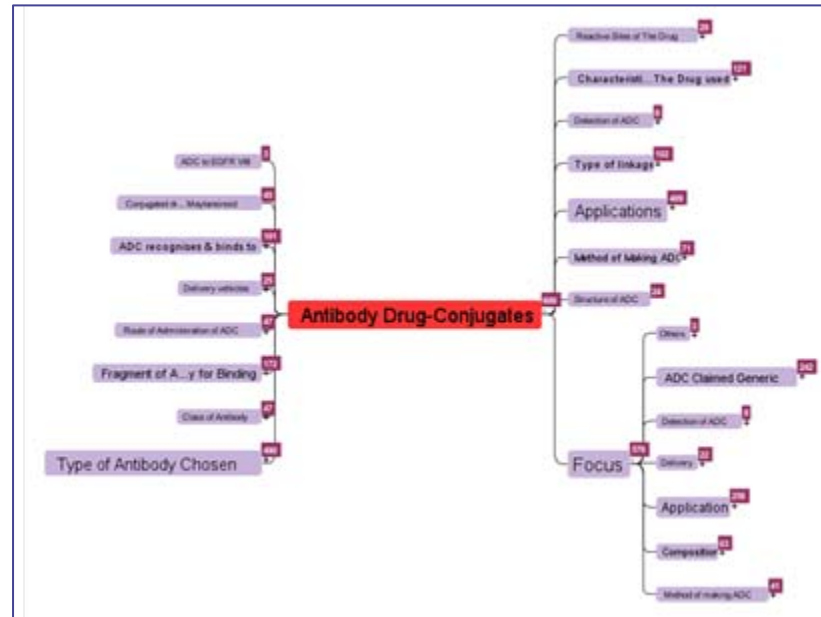
[From <http://www.biooncology.com>]

HUMAN-GENERATED PATENT LANDSCAPE (DOLCERA)



- PROS
 - Displayed on a web page as a FreeMind mind map
 - Leaf nodes linked to a Dashboard with access to individual documents
- CONS
 - Expensive and prone to errors and inconsistency
 - Little control over design – no update mechanism

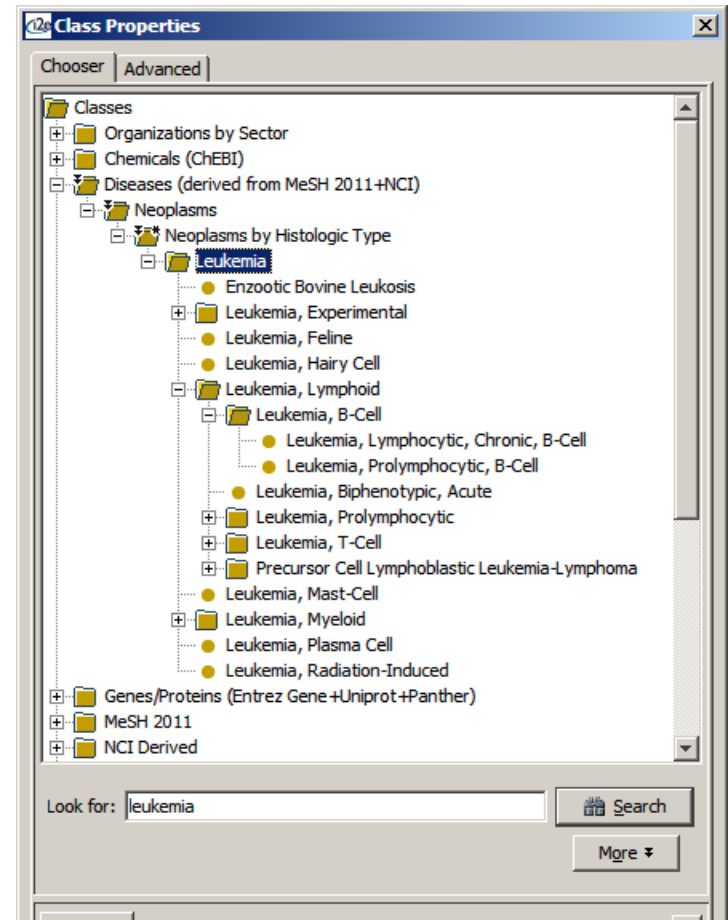
“COMPUTER-GENERATED” THESAURUS (INTELLIXIR)



- PROS
 - Displayed on a web page as an interactive applet
 - Leaf nodes linked to lists with access to individual documents
- CONS
 - Boolean queries not versatile enough for meaningful categorization

COMBINE HUMAN AND MACHINE STRENGTHS USING THE POWER OF I2E

- Machine: domain knowledge captured in I2E ontologies and macros; powerful, flexible queries
 - *Problem 1: How to display the results of class searches as interactive concept trees?*
- Human: expert review
 - *Problem 2: How to enable the user to rapidly select (and edit) the relevant assertions?*
 - Ontology shortcomings – for example, in Entrez Gene CS1 is a synonym for:
 - SLAMF7
 - CLSTN1
 - SSFA2
 - MYOZ2
 - CSH1
 - Patents are more difficult to mine than PubMed abstracts



PATENTS ARE NOT PUBMED: TYPOS, OCR ERRORS AND THE LEGAL LANGUAGE

WO1995007716: PHARMACEUTICAL COMPOSITION COMPRISING MONOCLONAL ANTIBODIES AGAINST THE INTERFERON RECEPTOR, WITH NEUTRALIZING ACTIVITY AGAINST TYPE I INTERFERON:

(WIPO) 5. Pharmaceutical composition according to anyone of **claims 1 to 4**, wherein the monoclonal antibody inhibits in vitro the binding of human type I-IFN, to the human cellular IFN-R when it is co-incubated with cells harboring the hu-IFN-R, **at a concentration of antibodies equal or inferior to 100 µg/ml, preferably equal or inferior to 50 µg/ml, advantageously inferior to 20 µg/ml, more preferably in the range of approximately 0,5 to 2 µg/ml.**

(PatBase) 5. Pharmaceutical composition according to anyone of **claims 1 to 4** wherein the monoclonal antibody inhibits in vitro the binding of human type I-IFN to the human cellular IFN-R when it is co-incubated with cells harboring the hu-IFN-R **at a concentration of antibodies equal or inferior to 100 g/ml preferably equal or inferior to 50 pg/ml advantageously inferior to 20 Zg/ml more preferably in the range of approximately 0 5 to 2 pg/ml.**

WO10048425A - COMPOSITIONS AND METHODS FOR TREATING THYMIC STROMAL LYMPHOPOIETIN (TSLP)-MEDIATED CONDITIONS

4. The method of claim 1, wherein the percentage of dissolved oxygen molecules present in the fluid as the charge- stabilized oxygen-containing nanostructures is a **percentage selected from the group consisting of greater than: 0.01 percent, 0.1 percent, 1 percent, 5 percent; 10 percent; 15 percent; 20 percent; 25 percent; 30 percent; 35 percent; 40 percent; 45 percent; 50 percent; 55 percent; 60 percent; 65 percent; 70 percent; 75 percent; 80 percent; 85 percent; 90 percent; and 95 percent.**

"When I use a word," Humpty Dumpty said in rather a scornful tone, "it means just what I choose it to mean -- neither more nor less."

WO2004032870 A2 (4/22/2004)

METHODS FOR TREATING POST-SURGICAL PAIN BY ADMINISTERING A NERVE GROWTH FACTOR ANTAGONIST AND COMPOSITIONS CONTAINING THE SAME

Abstract: The invention features methods and compositions for preventing or treating pain resulting from surgery or an incision by administering an antagonist of nerve growth factor (NGF). The NGF antagonist may be an anti-NGF (such as anti-hNGF) antibody that is capable of binding hNGF.

We claim:

1. A method for treating post-surgical pain in an individual comprising administering to the individual an effective amount of an antagonist of nerve growth factor (NGF), wherein the NGF antagonist is other than TrkA immunoadhesin.
2. The method of claim 1, wherein resting pain is suppressed or ameliorated.
3. The method of claim 1, wherein mechanically-induced pain is suppressed or ameliorated.
4. The method of claim 1, further wherein the NGF antagonist is other than an anti-NGF antibody.
5. The method of claim 1, wherein the NGF antagonist is a kinase inhibitor capable of inhibiting downstream kinase signaling associated with NGF receptor activity.
6. The method of claim 5, wherein the kinase inhibitor is K252a.
- ...
8. The kit of claim 7, further wherein the NGF antagonist is other than an anti-NGF antibody.

EP1556083 B1 (2/2/2011)

METHODS FOR TREATING POST-SURGICAL PAIN BY ADMINISTERING AN ANTIBODY AGAINST NERVE GROWTH FACTOR AND COMPOSITIONS CONTAINING THE SAME

Claims

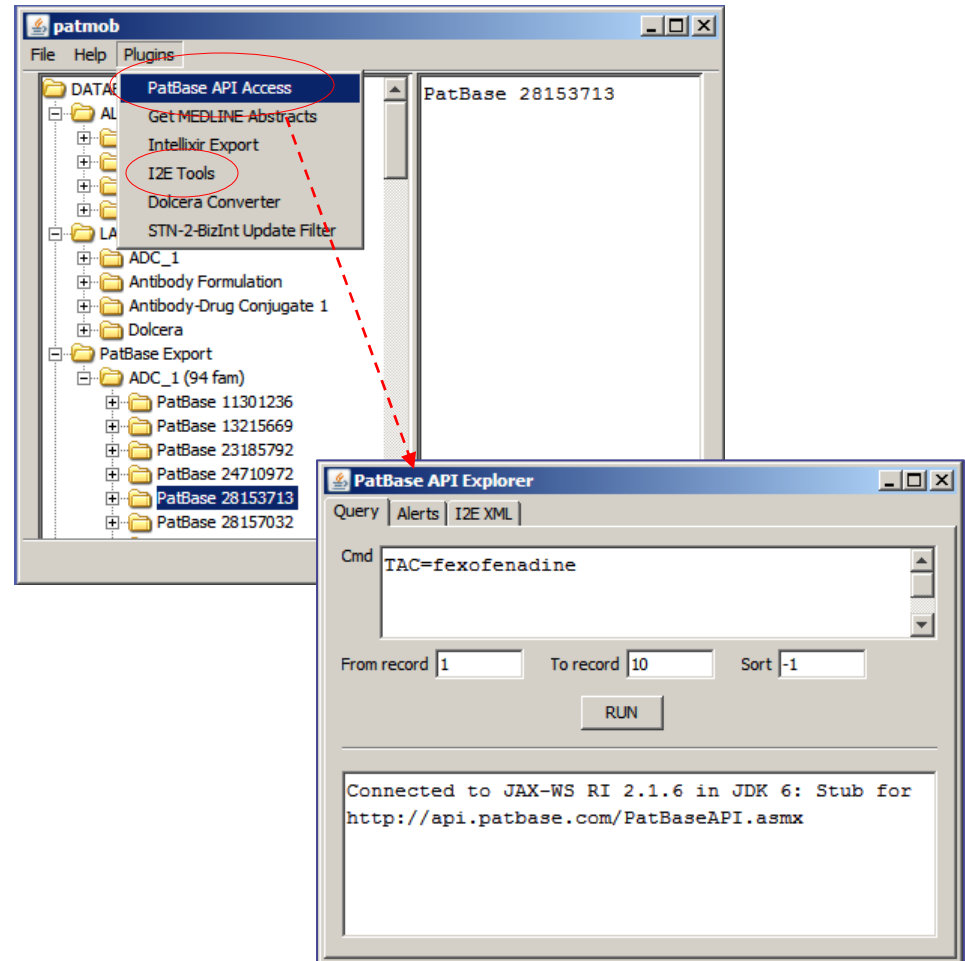
1. An anti-NGF antibody for use in the treatment of post-surgical pain.
2. An anti-NGF antibody of claim 1, wherein the anti-NGF antibody is for use in suppressing or ameliorating resting pain.
3. An anti-NGF antibody of claim 1, wherein the anti-NGF antibody is for use in suppressing or ameliorating mechanically-induced pain.
4. An anti-NGF antibody of any one of claims 1 to 3, wherein the anti-NGF antibody is a human antibody.
5. An anti-NGF antibody of any one of claims 1 to 3, wherein the anti-NGF antibody is a humanized antibody.
6. An anti-NGF antibody of any one of claims 1 to 3, wherein the anti-NGF antibody binds human NGF.
- ...
10. Use of an anti-NGF antibody in the manufacture of a medicament for treating post-surgical pain.
11. A pharmaceutical composition for use in the treatment of post-surgical pain comprising an anti-NGF antibody and a pharmaceutically acceptable carrier.

LANDSCAPING STRATEGY WORKFLOW

- Select a set of patent families in a broad search of a patent database
 - Working with a subset of relevant documents simplifies I2E queries
 - Patent families provide redundancy and verification
- Transfer the results to I2E for indexing
 - PatBase title, abstract and claims
- Repeatedly query the I2E index
 - Create queries to address different aspects of the landscape
 - Review and edit results in the context of patent families
 - Save relevant patent publications using the ontology/macro structure
- Combine all I2E searches
- Display the final result as a tree with links to documents

BRIEF INTRODUCTION TO PATMOB...

- A Java desktop application designed to facilitate rapid development of data convertors as plugins
- The infrastructure provides:
 - Relational database (JavaDB)
 - Internet access protocols
 - XML support
 - Tree editor
- PatBase API Access plugin
 - Connects to PatBase API
 - Allows querying
 - Downloads selected patent families
 - Converts PatBase XML to I2E-acceptable XML



...AND PATENT FAMILIES

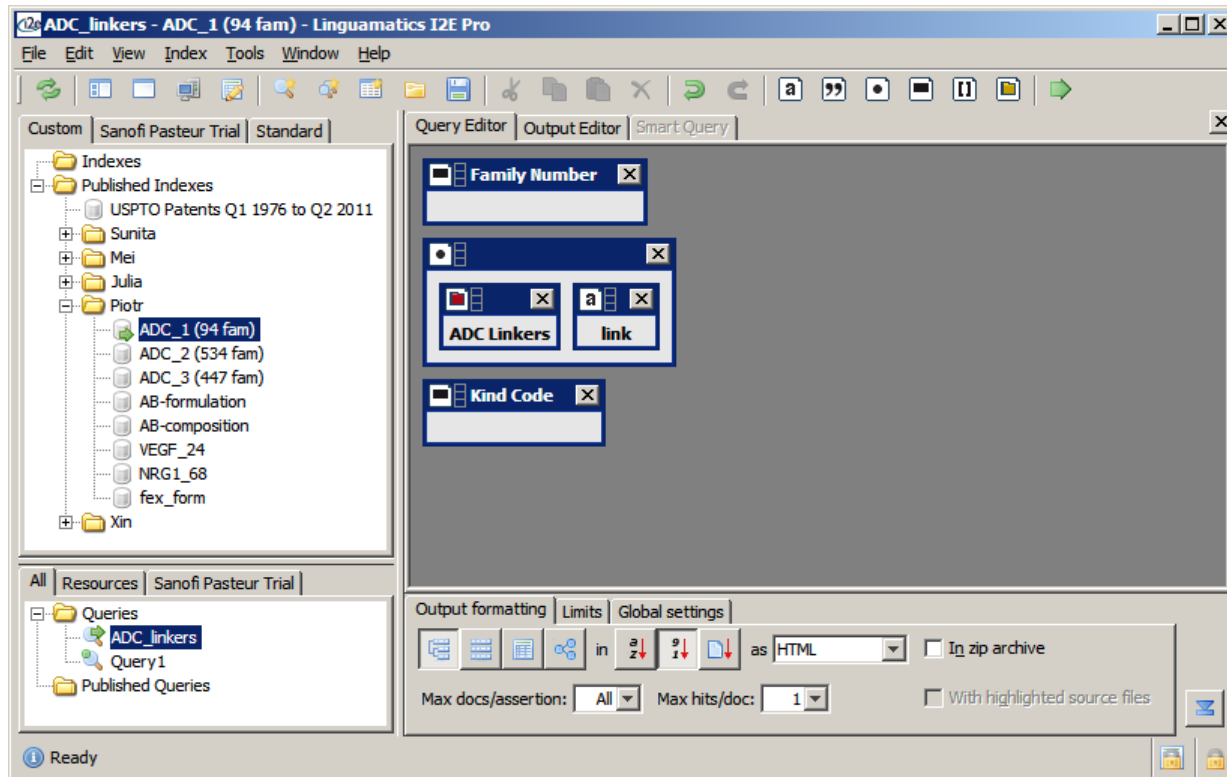
Real-time bibliographic data, with the full publication cycle of the patent application. Fetched from the Open Patent Services (OPS) at the European Patent Office.

The screenshot displays the 'Tree Branch Editor: PatBase Export' interface. On the left, a tree view shows a hierarchy of patent families under 'ADC_1 (94 fam)'. The selected entry is 'US2001036923 AA'. The main window shows the 'INFO' tab with the following bibliographic data:

```
US2001036923  Biblio:
=====
PN: US2001036923 A1
PD: 20011101
TI: Cytotoxic agents comprising modified doxorubicins and daunorubicins and their therapeutic use
PA: CHARI RAVI V.J, ; BLATTLER WALTER A, ; IMMUNOGEN INC;
AB: A cytotoxic agent comprising one or more modified doxorubicins/daunorubicin linked to a cell binding agent. A therapeutic composition for killing selected cell populations comprising: (A) a cytotoxic amount of one or more modified doxorubicins/daunorubicins covalently bonded to a cell binding agent through a linking group, and (B) a pharmaceutically acceptable carrier. A method for killing selected cell populations comprising contacting target cells or tissue containing target cells with an effective amount of a cytotoxic agent comprising one or more modified doxorubicins/daunorubicins linked to a cell binding agent. Novel sulfur-containing modified doxorubicins/daunorubicins.
=====
PN: US6630579 B2
PD: 20031007
TI: Cytotoxic agents comprising modified doxorubicins and daunorubicins and their therapeutic use
PA: CHARI RAVI V.J, ; BLATTLER WALTER A, ; IMMUNOGEN INC;
AB: A cytotoxic agent comprising one or more modified doxorubicins/daunorubicin linked to a cell binding agent. A therapeutic composition for killing selected cell populations comprising: (A) a cytotoxic amount of one or more modified doxorubicins/daunorubicins covalently bonded to a cell binding agent through a linking group, and (B) a pharmaceutically acceptable carrier. A method for killing selected cell populations comprising contacting target cells or tissue containing target cells with an effective amount of a cytotoxic agent comprising one or more modified doxorubicins/daunorubicins linked to a cell binding agent. Novel sulfur-containing modified doxorubicins/daunorubicins.
=====
```

At the bottom of the window, there are three buttons: 'Save to DB', 'Write to XML', and 'Write to TXT'.

LET'S I2E!



EXPORT I2E RESULTS AND OPEN WITH "I2E TOOLS" PLUGIN

The screenshot displays the Linguamatics I2E web interface. At the top, it shows search statistics: "Found 368 assertions from 987 hits in 147 docs. Examined all 465 docs. Took 0.11 secs (CPU 0.09). [more details]". The main area contains a table of results with columns for Family Number, ADC Linkers, link, Kind Code, Doc, and Hit. A "Select Columns for Parsing" dialog box is open in the foreground, showing a list of results columns and a "Columns for Categorization" section with fields for Fam (1 Family Number), Doc (7 Doc), KC (5 Kind Code), and Hit (9 Hit). In the top right of the interface, the "Redisplay as" dropdown is set to "XML". A red circle highlights this dropdown, and another red circle highlights the file name "ADC_linkers_00003171...xml" in the download bar at the bottom. A red arrow points from the "XML" dropdown to the file name.

[PT] Family Number	Family Number	ADC Linkers	link	[PT] Kind Code	Kind Code	Doc	Hit	
Family Number	24710972	Peptide	linkers	Kind Code	AA	7	US2003064984	224710972 Peptidyl prodrugs and linkers and stabilizers useful therefor ... AA
							US2006013860	224710972 Peptidyl prodrugs and linkers and stabilizers useful therefor ... AA
							US2003050331	124710972 ... AA... Also provided are peptidyl and disulfide linkers that are cleaved in vivo.
							US2003073852	124710972 ... AA... Also provided are peptidyl and disulfide linkers that are cleaved in vivo.
							US2005272798	124710972 ... AA... Also provided are peptidyl and disulfide linkers that are cleaved in vivo.
							US2009175888	124710972 ... AA... Also provided are peptidyl and disulfide linkers that are cleaved in vivo.
							US2009209734	124710972 ... AA... Also provided are peptidyl and disulfide linkers that are cleaved in vivo.
						6	US7087600	224710972 Peptidyl prodrugs and linkers and stabilizers useful therefor ... BB
					A1	2	EP2266986	124710972 ... A1... Also provided are peptidyl and disulfide linkers that are cleaved in vivo.
					E	1	USRE41252	224710972 Peptidyl prodrugs and linkers and stabilizers useful therefor ... E
					AA	7	US2003073852	224710972 Disulfide prodrugs and linkers and stabilizers useful therefor ... AA
					AA	6	US2006233794	1530868719 ... AA... 25. The method of claim 24 wherein the peptide linker is a dipeptide linker.
					AA	4	US2006074008	123185792 ... AA... Drug-Linker-Ligand Conjugates are disclosed in which a Drug is linked to a Ligand via a peptide-based Linker unit.
					AA	3	US2004202666	530774913 ... AA... The targeting antibody and the chemotherapeutic drug are linked via a linker comprising a hydrazide moiety.
					AA	2	US2008050310	444208713 ... AA... 84. The immun conjugate of claim 79 wherein the linker is attached to the antibody through a thiol group on the antibody.
					A2	1	WO07100385	136054046 ... A2... 25. The method of claim 23 wherein the linker reagent is a bis-maleimide reagent selected from DTME BMB BMD BMOE

EDITING I2E RESULTS

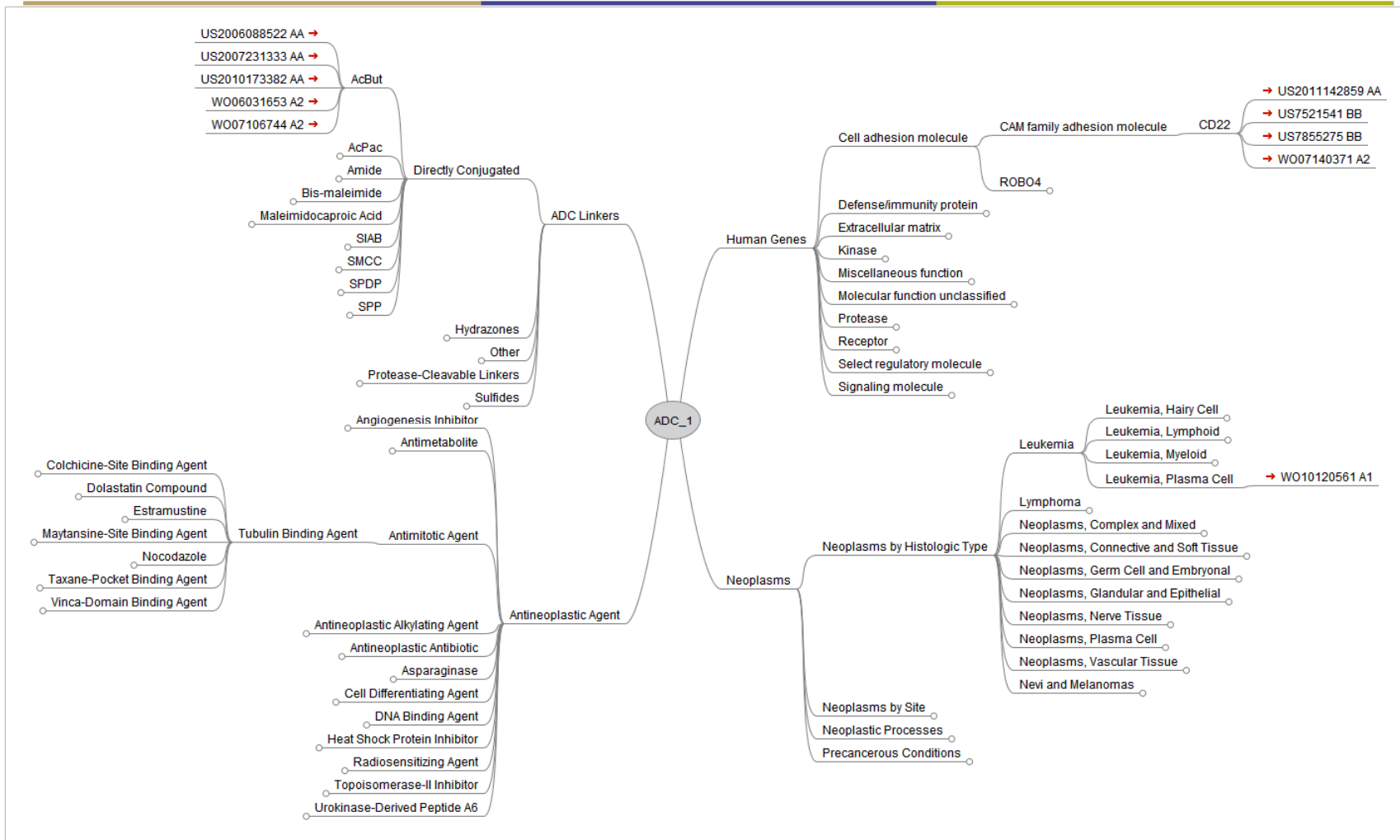
The image shows two overlapping windows. The background window is the 'Tree Branch Editor: I2E Results' application. It displays a hierarchical tree of patent nodes on the left. A red dashed arrow points from the 'US2003050331 AA' node in the tree to a context menu that is open over it. The menu options include 'Read at Patent Office', 'Read at PatBase Express', 'Get OPS Claims', 'Remove Node', 'Add Child Node...', and 'Set Hitte'. The 'Remove Node' option is highlighted. The foreground window is a web browser displaying the 'I2E Linguamatics I2E' interface. The address bar shows the URL 'https://www.i2eondemand.com/i2e/sanofi/server/tmp/ADC_linkers_00003687_111003_1604'. The page content includes a 'Family Number' field with the value '24710972', a 'Kind Code' field with the value 'BB', and a text area containing patent information. The text area includes the following text: '20061031 MEDAREX INC MEDAREX INC', 'The present invention provides analogues of duocarmycins that are potent cytotoxins.', and 'ADC Linkers link Also provided are peptidyl and disulfide linkers that are cleaved in vivo.' Below the text area, there is a paragraph of text: 'What is claimed is: <p>1. A compound having the structure: ##STR00034## or a pharmaceutically acceptable salt thereof, wherein X and Z are members independently selected from O S and NR 23 wherein R 23 is a member selected from H substituted or unsubstituted alkyl substituted or unsubstituted heteroalkyl and acyl; R 1 is H substituted or unsubstituted lower alkyl or C(O)R wherein R 8 is a member selected from NR 9R 10 and OR 9 in which R 9 and R 10 are members independently selected from H substituted or unsubstituted alkyl and substituted or unsubstituted heteroalkyl; R 2 is H or lower alkyl; R 3 is a member selected from (.dbd.O)'. The browser window also shows a 'Save to DB' button and a 'Write to X...' button at the bottom.

A TREE IN 3 STEPS

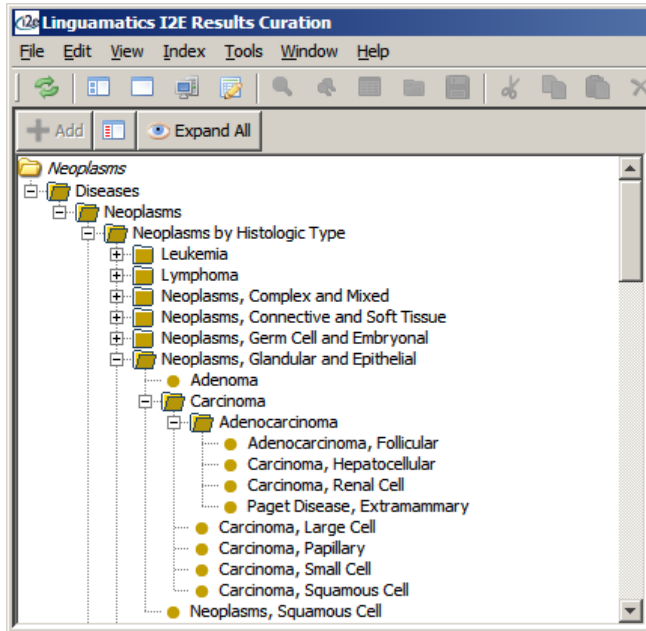
The image illustrates the process of building a hierarchical tree structure in three steps, shown in four sequential screenshots:

- Tree Branch Editor: I2E Results:** Shows a flat list of folders under the 'Tree' view, including 'I2E Results' and numerous 'PatBase' folders.
- Tree Branch Editor: Feature_Tree:** Shows the folders grouped into categories like 'AcBut', 'AcPac', 'Amide', 'BM(PEO)3', 'BM(PEO)4', 'IN07052DN2006 A', 'US2005276812 AA', 'BMB', 'BMDB', 'BMH', 'BMOE', 'DTME', 'SMCC', 'Thiol', 'BMB', 'BMDb', 'BMH', 'BMOE', 'Beta-glucuronide', 'DTME', 'Dipeptide', 'Disulfide', 'Hydrazide', 'Hydrazine', 'Hydrazone', 'MC-vc-PAB', 'Maleimide', 'Maleimidocaproic Acid', 'Malonate', 'PAB', 'Pentapeptide', 'Peptide', 'Phe-Lys', 'SIAB', 'SMCC', 'SPDP', 'SPP', 'Tetrapeptide', 'Thioether', 'Thiol', 'Tripeptide', and 'Val-Cit'.
- Tree Branch Editor: ADC Linkers:** Shows the folders grouped into categories like 'ADC Linkers', 'Directly Conjugated', 'AcBut', 'AcPac', 'Amide', 'Bis-maleimide', 'BM(PEO)3', 'BM(PEO)4', 'IN07052DN2006 A', 'US2005276812 AA', 'BMB', 'BMDb', 'BMH', 'BMOE', 'DTME', 'SMCC', 'Thiol', 'BMB', 'BMDb', 'BMH', 'BMOE', 'DTME', 'Maleimidocaproic Acid', 'SIAB', 'SMCC', 'SPDP', 'SPP', 'Hydrazones', 'Other', 'Protease-Cleavable Linkers', and 'Sulfides'.
- patmob:** Shows the final tree structure integrated into a larger application interface. The tree is displayed under 'DATABASE' and 'LANDSCAPING' with a search bar and various filters. A red arrow points from the 'ADC Linkers' folder in the previous step to the 'ADC Linkers' folder in this step.

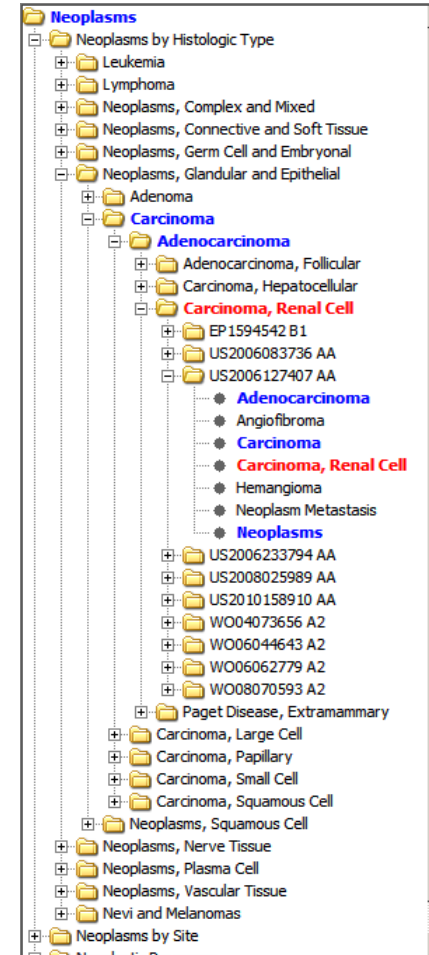
FREEMIND MIND MAP



ONTOLOGY CLASSES – SOME QUESTIONS



- Need for a simple I2E export of the ontology branches found in the search results.
- Displaying patents only in the leaf nodes gives cleaner results, but can miss important documents.



CONCLUSIONS

- I2E is a powerful tool for creating patent landscapes
- Feedback from experts/review articles to guide landscaping
- Best results obtained by a reiterative and recursive editing process
- Final product can be displayed in a variety of ways, but tools to manipulate the trees are still needed
 - E.g. list all documents shared by selected nodes

Thank you