Techniques for Evaluating Conflicting Information in Multi-Database Patents and Drug Pipeline Reports

ICIC 2010, Vienna

Diane Webb, President
27 October 2010

The Problem

- When creating pipeline and patents reports combining search results from different databases…
- *…how do you choose between conflicting information from different sources?*

The Problem – Patents example

	Title	Database	Patent Family			Abstract	Claims
			Patent	Kind	Date		
1	**Mouse trap used at home has enclosure which is provided with top and base having aperture and indentation that can be aligned to open enclosure for entry of mouse, such that contra-rotation of top relative to base is enabled to trap mouse.**	Derwent World Patents Index	WO 2005051079 EP 1691603	A1 A1	20050609 20060823	WO2005051079 A UPAB: 20050624 NOVELTY: The mouse trap has an enclosure having a top (1) and a base (3) respectively provided with an aperture (5) and an indentation (7). The manual rotation of the top relative to the base is enabled to open the enclosure with the alignment of the aperture and the indentation. The contra-rotation of the top relative to the base is performed upon entry of the mouse into the enclosure to trap the mouse. [..see record]	
2	**MOUSETRAP**	FAMPAT	GB 200325446 WO 200551079 AU 2004292376 EP 1691603 KR 20060110287 BR 200416076 US 20070017149 IN 2006CN01914 ZA 200603453 US 7506471 US 20090288332	D0 A1 A1 A1 A A A1 A A B2 A1	20031203 20050609 20050609 20060823 20061024 20070102 20070125 20070608 20070725 20090324 20091126	Disclosed is a mousetrap having an enclosure with a rotatable top part having a downwardly extending strike plate within the enclosure, and a bottom part having an upwardly extending catch plate within the enclosure, an aperture in each of the top part and the bottom part which are in substantial alignment when the mousetrap is set, and a trigger mechanism, wherein the mousetrap is set to incapacitate [..see record]	
3	**Mousetrap**	PatBase	GB 200325446 AU 2004292376 WO 05051079 EP 1691603 KR 2006110287 BR 200416076 US 2007017149	A0 AA A1 A1 A A A	2003-12-03 2005-06-09 2005-06-09 2006-08-23 2006-10-24 2007-01-02 2007-01-25	Source: WO05051079A1 The present invention is directed to a mousetrap comprising: an enclosure comprised of a top a base and apertures located on each of the top and the base wherein the enclosure is in an open position upon substantial alignment of the apertures; and a trigger mechanism comprising a lever arrangement and a biasing means operably connected to the top and the base wherein the lever [..see record]	
	MOUSETRAP	TotalPatent	WO 2005051079 AU 2004292376 BR PI0416076 BR PI0416076 EP 1691603	A1 A1 A1 A A1	2004-10-20	The present invention is directed to a mousetrap comprising: an enclosure comprised of a top, a base and apertures located on each of the top and the base, wherein the enclosure is in an open position upon substantial alignment of the	CLAIMS : 1. A mousetrap comprising: an enclosure comprised of a top, a base and apertures located on each of the top and the base, wherein the

BizInt Smart Charts 2010

The Problem – Drug Pipelines example

HER-2 Inhibitors (Apr 2008 updated with June 2008) - all dbs

	Product	Database	Highest Phase	Update Date	Mechanism of Action
12	**BMS-599626**	Thomson Reuters Integrity Compounds	Phase I		HER4 (erbB4) Inhibitors EGFR (HER1 erbB1) Inhibitors HER2 (erbB2) Inhibitors
13	**BMS-599626**	Citeline Pharmaprojects	No Development Reported	2007-01-22	ErbB-1 tyrosine kinase inhibitor (KI-TYE1-AN) ErbB-2 tyrosine kinase inhibitor (KI-TYE2-AN)
14	**AC-480**	Thomson Pharma	Discovery	2008-05-02	Anticancer EGFR family tyrosine kinase receptor inhibitor Erbb2 tyrosine kinase receptor inhibitor Erbb4 tyrosine kinase receptor inhibitor Epidermal growth factor antagonist
15	**BMS 599626**	Adis R&D Insight	Phase I	2008-05-08	Epidermal growth factor inhibitors HER2 inhibitors
16	**Pan-HER kinase inhibitor, Bristol-Myers Squibb**	IMS R&D Focus	Phase I	2007-01-01	EGF receptor inhibitor

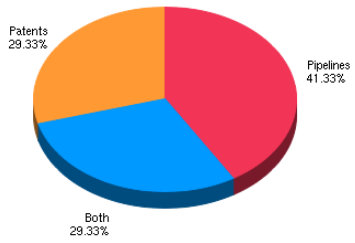
BizInt Smart Charts 2010

Survey conducted in September 2010

- SurveyGizmo survey with 13 questions – 6 patents related and 7 drug pipeline related
- Sent 300+ survey invitations to BizInt Smart Charts users; posted survey invitation on PIUG wiki and SLA DPHT list
- 75 complete responses

What kinds of reports do you create?

Answer	Count	Percent
Pipeline	31	41%
Patents	22	29%
Both	22	29%



Question 1: Patent Titles & Abstracts

- Derwent World Patents Index and Chemical Abstracts both provide enhanced titles and abstracts, while other databases, such as Patbase or MicroPatent, present the original or translated titles and abstracts.

When a record with enhanced titles and abstracts is part of a group, do you?

Question 1: Patent Titles & Abstracts

Answer	Count	Percent
b. select the title/abstract that best fits the search	19	45%
a. always select the enhanced title/abstract	11	26%
d. decide based on cost or license terms	9	21%
d. no preference	2	5%
c. never select the enhanced title/abstract	1	2%

Comments: Patent Titles & Abstracts

- "It all depends on the original query and sometimes I might even show both the original title and enhanced title."
- "We have a license with Patbase so that tends to influence the decision"
- "The title that best fit the search is very often the one from Derwent because more comprehensive."
- "The choice also depends on what the client wants to see."

Question 2: Consistency within a Record

- Selecting information from a mix of records can sometimes result in some fields not containing consistent information. For example, classification codes may come from a publication not in the selected family, or vice versa.

How important is consistency between fields?

Question 2: Consistency within a Record

Answer	Count	Percent
b. I would prefer that data between fields match, all else equal	22	52%
c. Absolute consistency betwen fields is not important	11	26%
a. I would always make sure the data between fields match	6	14%
d. no preference	3	7%

Question 3: Patent Families

- The families reported by databases are often different.

Which of these best describes how you select which publication numbers to show in your results?

Question 3: Patent Families

Answer	Count	Percent
b. Select the most comprehensive family	17	40%
c. Create a composite family combining all of the publications in all of the records	12	29%
a. Select the family from a particular database provider	10	24%
e. Select only publications for authorities of interest	2	5%
f. No Preference	1	2%

Comments: Patent Families

- "It depends on whether I am doing a state-of-the-art, patentability, FTO, or validity search."
- "I focus on US & PCT members when possible."
- "Authorities of interest are - for the time being - US, EP, WO."
- "It's important to be as comprehensive as possible for patent searches (and so in report too)."
- "It would be nice if there was an automated way to create a composite family."

Question 4: Publication Numbers

- If you select a basic publication number to represent the family, how do you choose?

BizInt Smart Charts 2010

Question 4: Publication Numbers

Answer	Count	Percent
a. Select the basic publication	14	34%
c. Select a specific authority	9	22%
b. Select the earliest publication in the retrieved results	6	15%
e. No preference	6	15%
d. Select an authority for which linked content is available	6	15%

BizInt Smart Charts 2010

Question 5: Claims - Language

- Assuming that you and/or your client are English-speakers -- If a patent has been granted, and the granted publication is currently only available in another language (e.g., German),

  Which claims would you present to your client?

Question 5: Claims - Language

Answer	Count	Percent
b. Machine translation of the granted claims	10	25%
c. Claims from any application in English	10	25%
a. The granted claims, even though they are not in English	9	22%
e. No preference	6	15%
d. English claims from an application to a specific authority	6	15%

Comments: Claims - Language

- "It would depend on what a particular attorney requested."
- "If the results are for a Freedom-To-Operate search, then the granted claims are most important. If the results are for a Prior Art or Landscape search, then the claims from an application in English are preferable."
- "US clients usually want to see US granted claims if available"

Question 6: Query Match

- Some search systems provide information about the quality of the match between a document and the query. One example is the percent identity or score from a BLAST alignment search.

When reporting results from such a system, do you…?

Question 6: Query Match

Answer	Count	Percent
d. No preference	14	38%
b. Summarize all queries and the quality of the match	13	35%
a. Display the best query match reported	9	24%
c. Display the retrieved document without reference to the match	1	3%

Applying the Rules

	Title	Database	Patent Family: Patent	Kind	Date	Abstract	Claims
1.	Mouse trap used at home has enclosure which is provided with top and base having aperture and indentation that can be aligned to open enclosure for entry of mouse, such that contra-rotation of top relative to base is enabled to trap mouse.	1.1 DWPI | link 1.2 TOTALPAT | link 1.3 Patbase | link 1.4 FAMPAT	GB 200325446	D0	20031203	WO2005051079 A UPAB: 20050624 NOVELTY: The mouse trap has an enclosure having a top (1) and a base (3) respectively provided with an aperture (5) and an indentation (7). The manual rotation of the top relative to the base is enabled to open the enclosure with the alignment of the aperture and the indentation. The contra-rotation of the top relative to the base is performed upon entry of the mouse into the enclosure to trap the mouse. [..see record]	CLAIMS : 1. A mousetrap comprising: an enclosure comprised of a top, a base and apertures located on each of the top and the base, wherein the enclosure is in an open position upon substantial alignment of the apertures;and a trigger mechanism comprising a lever arrangement and a biasing means operably connected to the top and the base, wherein the lever arrangement defines a gap through which a [..see record]
			WO 200551079	A1	20050609		
			AU 2004292376	A1	20050609		
			EP 1691603	A1	20060823		
			KR 20060110287	A	20061024		
			BR 200416076	A	20070102		
			US 20070017149	A1	20070125		
			IN 2006CN01914	A	20070608		
			ZA 200603453	A	20070725		
			US 7506471	B2	20090324		
			US 20090288332	A1	20091126		
	1.1 DWPI		1.4 FAMPAT			1.1 DWPI	1.2 TOTALPA
2.	Portable electrical trap for capturing and killing a mouse, has vacuum source which sucks the mouse fully into a collection chamber within which the mouse is subsequently suffocated.	2.1 DWPI | link 2.2 MPAT | link 2.3 TOTALPAT | link	US 6865843	B1	2005-03-15	US 6865843 B UPAB: 20050406 NOVELTY: Primary and secondary motion sensors (28,34) detect the presence of a mouse inside the interior cavity of the mouse trap (10). A primary gate and a secondary gate (36) in turn automatically opens upon activation of the corresponding motion sensor. A vacuum source (40) sucks the mouse fully into a collection chamber (38) within which the mouse is subsequently suffocated. [..see record]	1. A portable electrical mouse trap for capturing and killing a mouse, comprising: a housing in the shape of a cat having a head with a pair of eyes and a mouth defining an interior cavity, the mouth forming an entrance to the interior cavity; a primary motion sensor near the entrance to the interior cavity for selectively closing the mouth; a retractable primary gate in electrical

Question 1: Criteria for Selecting Records

- Searching multiple databases often retrieves a product from more than one database, as shown in the example below.

  If you were to select a single record to represent this compound, what would be the most important feature influencing your choice?

Question 1: Criteria for Selecting Records

Answer	Count	Percent
Select the record with the most recent record update date	20	37%
Select the record with information included in most all of the cells	18	33%
Select the record with the Highest Phase?	7	13%
Select a specific database above all others if it is present?	5	9%
No preference	4	7%

Question 2: Preferred Database

- Everything else being equal, in what order would you choose to take information from the pipeline databases? (1 = 1st choice, 2 = 2nd choice, etc.)
  - Thomson Pharma/IDdb3
  - Thomson Prous Integrity
  - IMS R&D Focus
  - Adis R&D Insight
  - Citeline Pipeline/Pharmaprojects
  - other

Question 2: Preferred Database

Database	1st	2nd	3rd	4th	5th	6th	Avg Rank
Thomson Pharma/IDdb	27	15	8	1	1		1.7
Adis R&D Insight	16	9	16	10	2		2.5
Citeline Pipeline/ Pharmaprojects	8	15	14	8	2	2	2.7
Thomson Prous Integrity	3	9	5	12	18	2	3.8
IMS R&D Focus		4	8	17	16		4.0
Other		1		1	4	21	5.6

Comments: Preferred Database

- "For very early stage compounds, I lean toward Integrity, whereas for compounds in clinical trials I rely on IDdb or Adis."
- "I prefer databases where I know that the end-user has a license ."
- "Adis just seems to have the most recent information and the most flushed out records."
- "I like the commercial summaries of Citeline Pipeline and IMS R&D Focus; they consistently cite their information sources."

Question 3: Mechanism of Action

- In this example, if your goal is to identify compounds in development for the target "HER2", how would you select the Mechanism of Action to appear in the reference row?

	Product	Common Drug Name	Database	Originator	Mechanism of Action	Last Update
383	**TAK 285**	TAK-285	IMS R&D Focus	Takeda (Japan)	protein kinase inhibitor tyrosine kinase inhibitor signal transduction inhibitor	2009-08-24
384	**TAK-285**	TAK-285	Thomson Pharma	Takeda Pharmaceutical Co Ltd	Anticancer Erbb2 tyrosine kinase receptor inhibitor Anticancer protein kinase inhibitor AMP activated protein kinase stimulator	2010-04-02
385	**TAK 285**	TAK-285	Adis R&D Insight	Takeda (Originator)	Epidermal growth factor receptor antagonists HER2 inhibitors	2010-02-16
386	**TAK-285**	TAK-285	Thomson Reuters Integrity Compounds	Millennium Pharmaceuticals Takeda	EGFR (HER1 erbB1) Inhibitors HER2 (erbB2) Inhibitors	
387	**TAK-285**	TAK-285	Citeline Pipeline	Takeda	ErbB-1 tyrosine kinase inhibitor ErbB-2 tyrosine kinase inhibitor	2010-02-05

Question 3: Mechanism of Action

Answer	Count	Percent
Select the cell that has HER2 listed even if it is not the only mechanism listed	27	50%
Select the cell with the largest list of Mechanisms of Action	7	13%
No preference	6	11%
Prefer a specific database regardless of what Mechanism of Actions are listed	6	11%
Select the cell that has HER2 listed as the only Mechanism of Action	4	7%
Select the Mechanism of Action cell from the most recently updated record	4	7%

Comments: Mechanism of Action

- "If I'm looking for a specific MOA, I want to see that spelled out."
- We create our own column for MoA, so we have a standard set that we pull from, but still include the list from the database.
- After selecting the cell with the most MOAs, I would cut and paste any I missed from the other records (if they are correct .)
- I'd select the entry with the most accurate details.

Question 4: Originator vs. Licensees

- As companies merge or are acquired, some database producers re-assign the most current company name to the "Originator" company field. When selecting a cell to represent the Originator company name, do you?

	Product	Database	Synonyms	Originator	Licensee	Last Update
225	**Neratinib**	Adis R&D Insight	HKI 272 HKI-272	Wyeth (Originator)		2010-03-19
226	**neratinib**	Thomson Pharma	HKI-272	Wyeth	Pfizer Inc	2010-03-19
227	**neratinib (pINN)**	IMS R&D Focus	neratinib HKI 272	Pfizer		2010-03-10
228	**neratinib**	Citeline Pipeline	CDK4 kinase inhibitor, Wyeth HKI-272 WAY-179272	Pfizer		2010-01-08
229	**Neratinib**	Thomson Reuters Integrity Compounds	HKI-272 WAY-179272-B (maleate) WAY-179272	Wyeth Pharmaceuticals Pfizer		

Question 4: Originator vs. Licensees

Answer	Count	Percent
Select the current Company actively developing the product, even if it was not the original company.	19	36%
Not worry about which company is listed as originator	18	34%
Select the "Licensee" field if it contains the "actual" originating company, since that company holds the patents and other intellectual property rights for the product	16	30%

Question 5: Phase of Development

- Many times there are multiple indications in development for a given product and these indications are frequently in different phases of development. When you are interested in one particular indication, what would you choose for the Highest Phase?

	Product	Database	Originator	Highest Phase	Drug Development Phase			Last Update
					Indication	Phase	Country	
12	**AC-480**	Thomson Pharma	Bristol-Myers Squibb Co	Phase 2 Clinical	Cancer Solid tumor	Discontinued Phase 2 Clinical	US US	2009-12-14
13	**AC-480**	Citeline Pipeline	Ambit Biosciences	Phase II	Cancer, general Cancer, brain	Phase II Phase I		2009-11-26
14	**AC 480**	IMS R&D Focus	Bristol-Myers Squibb, USA	Phase II	NSCLC	Phase II	USA	2009-07-06
15	**BMS 599626**	Adis R&D Insight	Bristol-Myers Squibb (Originator)	Phase II	Solid tumours	Phase II	USA	2009-05-06
16	**BMS-599626**	Citeline Pipeline	Bristol-Myers Squibb	No Development Reported	Cancer, breast	No Development Reported		2007-01-22
17	**AC-480 (free base)**	Thomson Reuters Integrity Compounds	Bristol-Myers Squibb	Phase I	Cancer, solid tumor	Phase I	-	

Question 5: Phase of Development

Answer	Count	Percent
The highest phase for the specific indication I'm interested in	36	67%
The highest phase listed for any indication	11	20%
The phase from the most recently updated record	5	9%
No preference	2	4%

Question 6: Highest Phase for Indication

- Launched products frequently also have multiple indications at lower phases of development. In the example, the Highest Phase for Lapatinib (Tykerb) is “Launched”.

If you are asked to prepare a summary of drugs in development for the treatment of brain cancer what would you want in the “Highest Phase” or other sortable phase column?

Question 6: Highest Phase for Indication

Highest Phase	Drug Development Phase	
	Indication	Phase
Launched	Cancer, breast	Launched
	Cancer, lung, general	Phase III
	Cancer, head and neck	Phase III
	Cancer, renal	Phase III
	Cancer, bladder	Phase III
	Cancer, gastrointestinal, stomach	Phase III
	Cancer, oesophageal	Phase III
	Cancer, brain	Phase II
	Cancer, ovarian	Phase II
	Cancer, peritoneal	Phase II
Launched	Breast cancer (Combination therapy, Second-line therapy)	Launched
	Brain cancer (Metastatic disease:Brain metastases arising from breast cancer)	Phase III
	Breast cancer (Combination therapy, First-line therapy)	Phase III
	Breast cancer (Adjunctive treatment)	Phase III
	Gastric cancer (Combination therapy)	Phase III
	Head and neck cancer	Phase III
	Solid tumours	Phase II

Question 6: Highest Phase for Indication

Answer	Count	Percent
The phase from the most recently updated record	37	69%
The phase of the indication I am interested in	10	19%
The "highest" phase of development listed for any indication	5	9%
No preference	2	4%

Comments: Highest Phase for Indication

- "A separate and sortable phase column that is tied to the indication is the most useful."
- "It is important to note that the drug has been launched for another indication."
- "I like to keep track of both the highest indication and the phase the product is in for a given indication."

Question 7: Key Information

What is the most important piece of information to consider when choosing which drug records to display in your report?

Listed Product Name, Database, Synonyms, Originating Company, Indication, Update Date, Mechanism of Action, Highest Phase Listed

1 = first choice, 2 = second choice, etc.)

Question 7: Key Information

Database	Average Rank
Product Name Listed (31 chose as “1”)	2.6
Indications	3.8
Mechanism of Action	4.1
Highest Phase	4.2
Synonyms	4.7
Originating Company	4.9
Update Date	5.3
Database	6.5

Applying the Rules

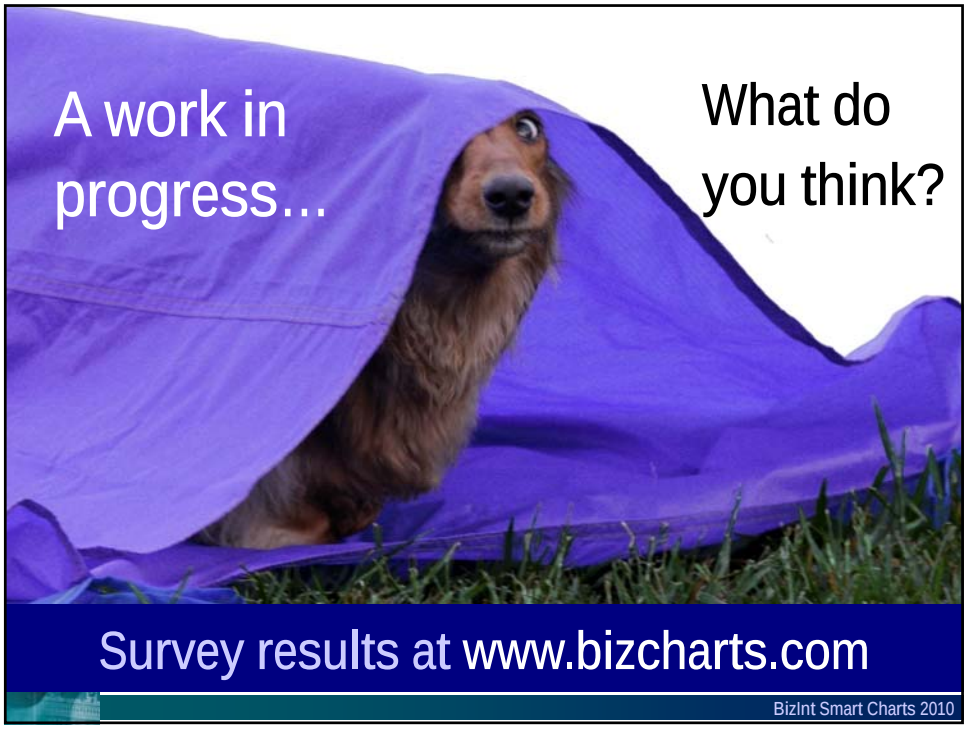
	Product	Database	Highest Phase	Update Date	Mechanism of Action
3.1	AC-480 ✔	Thomson Pharma	Discovery	2008-05-02	Anticancer ✔ EGFR family tyrosine kinase receptor inhibitor Erbb2 tyrosine kinase receptor inhibitor Erbb4 tyrosine kinase receptor inhibitor Epidermal growth factor antagonist
3.2	BMS- 6	Adis R&D Insight	Phase I ✔	2008-05-08 ✔	Epidermal gro or inhibitors HER2 inhibitors
3.3	BMS- 6	Thomson Reuters Integrity Compounds	Phase I		HER4 (erbB4) I rs EGFR (HER1 erbB1) Inhibitor HER2 (erbB2) Inhibitors
3.4	Pan-HER kinase inhibitor, Bristol-Myers Squibb	IMS R&D Focus	Phase I	200 1	EGF receptor inhibitor
3.5	BMS-599626	Citeline Pharmaprojects	No Development Reported	2007-01-22	ErbB-1 tyrosine kinase inhibitor (KI-TYE1-AN) ErbB-2 tyrosine kinase inhibitor (KI-TYE2-AN)

BizInt Smart Charts 2010

Applying the Rules

	Product	Database	Highest Phase	Update Date	Mechanism of Action
1.	ARRY-380	1.1 TPharm | link 1.2 Adis | link 1.3 Integr | link 1.4 Integr | link 1.5 PProj 1.6 PProj	Phase I Clinical Trial	2008-05-08	Anticancer Erbb2 tyrosine kinase receptor inhibitor
	1.1 TPharm		1.5 PProj	1.5 PProj	1.1 TPharm
2.	BIBW-2992	2.1 TPharm | link 2.2 Adis | link 2.3 Integr | link 2.4 IMS 2.5 PProj	Phase II/III	2008-06-10	Anticancer EGFR family tyrosine kinase receptor inhibitor Erbb2 tyrosine kinase receptor inhibitor
	2.1 TPharm		2.2 Adis	2.2 Adis	2.1 TPharm
3.	AC-480	3.1 TPharm | link 3.2 Adis | link 3.3 Integr | link 3.4 IMS 3.5 PProj 3.6 PProj	Phase I	2008-05-08	Anticancer EGFR family tyrosine kinase receptor inhibitor Erbb2 tyrosine kinase receptor inhibitor Erbb4 tyrosine kinase receptor inhibitor Epidermal growth factor antagonist
	3.1 TPharm		3.2 Adis	3.2 Adis	3.1 TPharm

Charts 2010


A work in progress…
What do you think?
Survey results at www.bizcharts.com
BizInt Smart Charts 2010