Software for Business Intelligence

BizInt Smart Charts

Enhancing Patent Family Display in BizInt Smart Charts in Patents

PIUG 2015 Annual Conference, Lombard IL

May 2015

John Willmore, VP Product Development

Searching multiple patent databases

- Searching multiple databases is a common strategy to increase recall, whether due to indexing, coverage, or timeliness.
- Crossing results to a single platform is a common approach, but one that **discards the value added content of the other sources**.
- But, presenting results from these different platforms in one report can be challenging.

GenomeQuest + LifeQuest

- GenomeQuest searches biological sequences found in patent publications
- LifeQuest uses advanced indexing to search life sciences patents
- Set management features in LifeQuest allow you to tag (color) publications identified by the sequence searches

LifeQuest color tagging in BizInt Smart Charts

- GenomeQuest sequence results transferred to LifeQuest are “color coded” in new set.

LifeQuest: Arachis

	Title	Patent Family			Priority Date	Color	Pub. Status	Abstract	Inventor(s)
		Patent	Kind	Date					
7	**Vaccine carrier**	AU2007257308	B2	2013-06-06	2007-06-11	3	Grant	(57) Abstract: The present invention relates to a hypoallergenic protein consisting of at least one hypoallergenic molecule derived from an allergen, which is fused or conjugated to at least one second non-allergenic protein or fragment thereof.	Gronlund, Hans Focke-Tejkl, Margarete Valent, Peter Tinhofer, Johanna Reininger, Renate Popow-Kraupp, Theresia Valenta, Rudolf Vrtala, Susanne Westritschnig, Kerstin Van Hage, Marianne Spitzauer, Susanne Linhart, Birgit Swoboda, Ines
8	**Pharmaceutical formulations and the use thereof for the treatment of peanut allergy**	AU2012351541	A1	2014-07-03	2012-12-14	0	Application	The present invention relates to compositions which can be used in immunotherapy and especially to compositions which can be used in immunotherapy for mammals, such as human mammals, suffering from peanut allergy. The present invention further relates to the use of the present compositions for the therapeutic treatment for desentizing the immune system of a mammal suffering from an	Koppelman, Stefan Johan Van Der Kleij, Joanna Paulina Maria

The value of preserving the original hit

- When you only transfer PNs to the final result, the details of the query are no longer available for your report. The new context is a number search.
- Stitching data from the original queries into your result builds a richer report, enables analysis that can't be performed on a single platform.

GenomeQuest + LifeQuest (2)

- Including information from the subject sequence improves the value to the end user.

	Title	Database	Patent Family			Pub. Status	Color	Patent Sequence Location	Inventor(s)
			Patent	Kind	Date				
1.	Method and Kit for the detection of allergens	1.1 LIFEQ | *link* 1.2 GPATNUC | *link*	EP2226395	A1	2010-09-08	Application	2	claim: 1; 3; 4; 6; 7; 14	Marmiroli, Nelson Gulli', Mariolina Pafundo, Simona
	1.1 LIFEQ				1.1 LIFEQ	1.1 LIFEQ	1.1 LIFEQ	1. NUC	
2.	PREVENTION AND TREATMENT OF NOSEMA DISEASE IN BEES	2.1 LIFEQ | *link* 2.2 GPATNUC | *link*	US20140371298	A1	2014-12-18	Application	2	claim: 12; 34	Nitzan PALDI Eitan Glick
	2.1 LIFEQ				2.1 LIFEQ	2.1 LIFEQ	2.1 LIFEQ	2.2 GPATNUC	
3.	Methods of genetic analysis of mouse	3.1 LIFEQ | *link* 3.2 LIFEQ | *link* 3.3 GPATNUC | *link* 3.4 GPATNUC | *link*	US7250289	B2	2007-07-31	Grant	2	claim: 1; 2; 3; 5	Xue Mei Zhou

Using the PN list to get new content

- Transfer the PN list from your final result set to a different platform to get unique content.
- In this example, we retrieve the corresponding families from Orbit, and add the Family Legal Status to our report.
- LifeQuest publications are grouped using FAMPAT extended families (based on Common Patent Family).

LifeQuest + Orbit (1)

- Revealing interesting status details for the retrieved documents…

	Title	Database	Document Strategy: Pub. Number	Color		Priority Date	Color	Family Status: Pub No.	State	Status	Expiry	Pub. Status
4.	Packaged virus-like particles for use as adjuvants: method of preparation and use	14.1 LIFEQ | link 14.2 LIFEQ | link 14.3 FAMPAT | link	CN1662253A AU2003242742B2 WO200400351A1	5 0	14.1 14.2 14.3	2002-06-20	5	EP1513552 DE60335186 HK1074578 WO2004000351 AT489969 AU2003242742 BR0311995 CA2488856	ALIVE DEAD ALIVE ALIVE DEAD DEAD DEAD DEAD	GRANTED LAPSED GRANTED PENDING LAPSED EXPIRED LAPSED LAPSED	2023-06-20 2013-01-01 2023-06-20 2023-06-20 2011-06-15 2013-01-17 2013-04-16 2012-05-08	Grant

AU2003242742 DEAD EXPIRED 2013-01-17

								JP2006502979 NZ537002 PL375306 RU2005101206 ZA200408709 US2004005338 US2011070267	DEAD DEAD DEAD DEAD ALIVE DEAD ALIVE	LAPSED LAPSED REVOKED REVOKED GRANTED LAPSED PENDING	2013-10-01 2014-01-31 2012-12-31 2013-04-20 2023-06-20 2009-01-08 2023-06-20	
	14.1 LIFEQ					14.3 FAMPAT	14.1 LIFEQ				14.3 FAMPAT	14.2 LIFEQ

LifeQuest + Orbit (2)

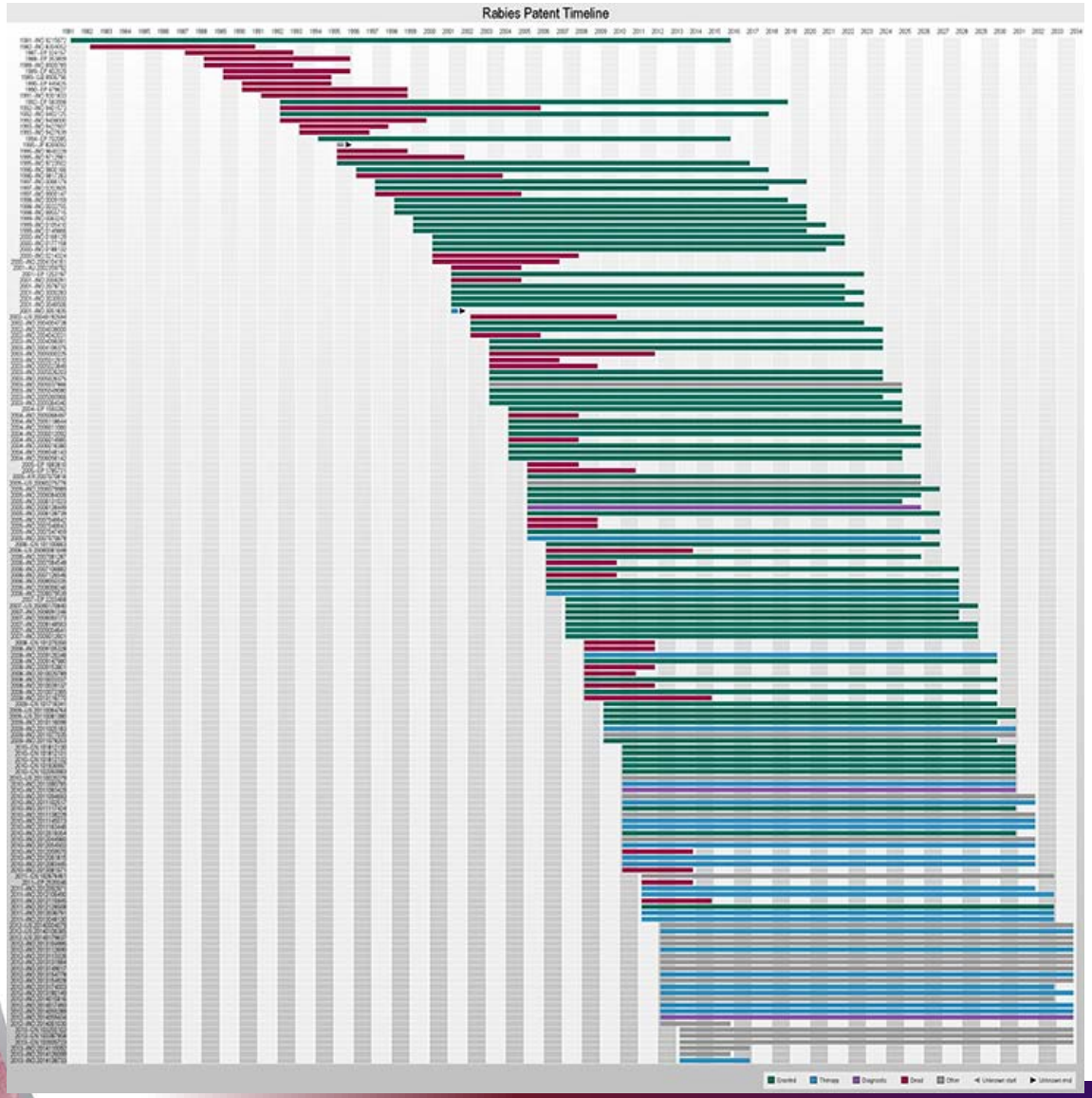
- And summarizing the search strategy (color) for each retrieved document in the family.

Title	D	Document Strategy		Priority	Color	Family Status				Pub. Status
						Pub No.	State	Status	Expiry	
4. Packaged virus-like particles for use as adjuvants: method of preparation and use	1	CN1662253A	5	14.1	-20	5	EP1513552	ALIVE	GRANTED	2023-06-20	Grant
	1	AU2003242742B2	0	14.2			DE60335186	DEAD	LAPSED	2013-01-01	
	1	WO200400351A1		14.3			HK1074578	ALIVE	GRANTED	2023-06-20	
							WO2004000351	ALIVE	PENDING	2023-06-20	
							AT489969	DEAD	LAPSED	2011-06-15	
							AU2003242742	DEAD	EXPIRED	2013-01-17	
							BR0311995	DEAD	LAPSED	2013-04-16	
							CA2488856	DEAD	LAPSED	2012-05-08	
							CN1662253	DEAD	LAPSED	2012-11-21	
							IL164812	DEAD	LAPSED	2010-12-18	
							IN241952	ALIVE	GRANTED	2023-06-20	
							JP2006502979	DEAD	LAPSED	2013-10-01	
							NZ537002	DEAD	LAPSED	2014-01-31	
							PL375306	DEAD	REVOKED	2012-12-31	
							RU2005101206	DEAD	REVOKED	2013-04-20	
							ZA200408709	ALIVE	GRANTED	2023-06-20	
							US2004005338	DEAD	LAPSED	2009-01-08	
							US2011070267	ALIVE	PENDING	2023-06-20	
	14.1 LIFEQ			14.3 FAMPAT	14.1 LIFEQ				14.3 FAMPAT	14.2 LIFEQ	

Using the PN list to get new content (again)

- Once you have your final result set, you may still want to retrieve the corresponding records from a new system for visualization.
- In this example, we used domain specific indexing in Thomson Reuters Cortellis to find our result set.
- The corresponding families from Orbit provided legal status and calculated expiry dates for the visualization.

Visualization of Integrated Data



Orbit:
Family State
Calculated Expiry

Thomson Cortellis:
Patent Type
Product Type

Integrating unique data content

- Once you have your final result set, you may still want to retrieve the corresponding records from a new system for visualization.
- In this example, we used domain specific indexing in Thomson Reuters Cortellis to find our result set.
- The corresponding families from Orbit provided legal status and calculated expiry dates for the visualization.

Integrating unique data content

- Thomson Reuters Cortellis indexes pharmaceutically relevant patents.

Enhanced Title	Indications	Patent Type	Classifications
Ebola virion proteins which are useful for the detection of and vaccination against ebola virus infection.	Ebola virus infection		Anti-Infectives Biologicals and Immunologicals
Monoclonal antibodies and vaccines against epitopes on the Ebola virus glycoprotein	Ebola virus infection	Product	Anti-Infectives Biologicals and Immunologicals

Integrating unique data content

- Questel FAMPAT provides family legal status data.

Family Status			
Pub No.	**State**	**Status**	**Expiry**
WO2012050193	DEAD	LAPSED	2013-12-03
JP2014005205	ALIVE	PENDING	2030-10-14
WO2011071574	ALIVE	PENDING	2030-09-01
EP2473525	DEAD	LAPSED	2014-08-27
US2012164153	ALIVE	PENDING	2030-09-01

Integrating unique data fields

- Records are matched based on publication numbers in common ("Common Family")

Enhanced Title	Indications	Patent Type	Classifications	Family Status			
				Pub No.	State	Status	Expiry
Monoclonal antibodies and vaccines against epitopes on the Ebola virus glycoprotein	Ebola virus infection	Product	Anti-Infectives Biologicals and Immunologicals				
				WO200116183	DEAD	LAPSED	2006-03-26
				AU7089600	DEAD	LAPSED	2006-03-26
				US6630144	ALIVE	GRANTED	2020-08-29
Monoclonal antibodies against glycoprotein of Ebola Sudan Boniface (ESB) virus - useful in the diagnosis and treatment of ESB virus infection.	Ebola virus infection	Diagnostic, Analysis and Assay Product (Macromolecule)	Anti-Infectives Biologicals and Immunologicals Diagnostics				
				WO2011071574	ALIVE	PENDING	2030-09-01
				EP2473525	DEAD	LAPSED	2014-08-27
				US2012164153	ALIVE	PENDING	2030-09-01
Ebola virus liposome vaccines - useful in eliciting immune responses against Ebola virus infection.	Ebola virus infection	Formulation	Anti-Infectives Biologicals and Immunologicals Pharmaceutics				
				WO2012050193	DEAD	LAPSED	2013-12-03
				JP2014005205	ALIVE	PENDING	2030-10-14

Integrating unique data content

- And fused into a single virtual row based on selection rules in Reference Rows.

	Enhanced Title	Indications	Patent Type	Classifications	Family Status				Database
					Pub No.	State	Status	Expiry	
2.	**Monoclonal antibodies and vaccines against epitopes on the Ebola virus glycoprotein**	Ebola virus infection	Product	Anti-Infectives Biologicals and Immunologicals	WO200116183 AU7089600 US6630144	DEAD DEAD ALIVE	LAPSED LAPSED GRANTED	2006-03-26 2006-03-26 2020-08-29	2.1 CORTP | *link* 2.2 FAMPAT | *link*
	2.1 CORTP	2.1 CORTP	2.1 CORTP	2.1 CORTP				2.2 FAMPAT	
3.	**Monoclonal antibodies against glycoprotein of Ebola Sudan Boniface (ESB) virus - useful in the diagnosis and treatment of ESB virus infection.**	Ebola virus infection	Diagnostic, Analysis and Assay Product (Macromolecule)	Anti-Infectives Biologicals and Immunologicals Diagnostics	WO2011071574 EP2473525 US2012164153	ALIVE DEAD ALIVE	PENDING LAPSED PENDING	2030-09-01 2014-08-27 2030-09-01	3.1 CORTP | *link* 3.2 FAMPAT | *link*
	3.1 CORTP	3.1 CORTP	3.1 CORTP	3.1 CORTP				3.2 FAMPAT	
4.	**Ebola virus liposome vaccines - useful in eliciting immune responses against Ebola virus infection.**	Ebola virus infection	Formulation	Anti-Infectives Biologicals and Immunologicals Pharmaceutics	WO2012050193 JP2014005205	DEAD ALIVE	LAPSED PENDING	2013-12-03 2030-10-14	4.1 CORTP | *link* 4.2 FAMPAT | *link*
	4.1 CORTP	4.1 CORTP	4.1 CORTP	4.1 CORTP				4.2 FAMPAT	
5.	**Chimeric filovirus glycoproteins useful in vaccines against Ebola and Marburg virus infections**	Marburg virus infection Ebola virus infection	Product	Anti-Infectives Biologicals and Immunologicals	WO02079239 US7731975	DEAD DEAD	LAPSED LAPSED	2006-03-29 2014-06-08	5.1 CORTP | *link* 5.2 FAMPAT | *link*
	5.1 CORTP	5.1 CORTP	5.1 CORTP	5.1 CORTP				5.2 FAMPAT	

Integrating data from a single query

- Certain databases return results at a finer level of detail than you want in a final report
- Landscapes from a publication level database can provide more insight at the family level.
- Similar integration techniques can be used in these cases.

Summarize TotalPatent results by family

- Example: reduce publications to a single family display from Lexis Nexis TotalPatent results

Title	Patent Family			Abstract	Claims
	Patent	Kind	Date		
HELMET	DE 50303117	D1	2006-06-01	The invention relates to a helmet (12), particularly a bicycle helmet, comprising an adjusting device (30, 40) in order to adapt the size of the helmet (12) to the size of the head of an individual wearing the helmet (12). A rear light (60) is provided on the adjusting device (30, 40).	
	AT 324054	T			
	EP 1513424	A2			
	EP 1513424	B1			
	WO 2004000054	A2			
	WO 2004000054	A3			
	WO 2004000054	A8			
HELMET	EP 1513424	B1	2006-04-26	The invention relates to a helmet (12), particularly a bicycle helmet, comprising an adjusting device (30, 40) in order to adapt the size of the helmet (12) to the size of the head of an individual wearing the helmet (12). A rear light (60) is provided on the adjusting device (30, 40).	Helmet (12, 12.2), particularly a bicycle helmet, comprising an adjusting device for adapting the size of the helmet to the head size of a person wearing the helmet, characterized in that- fastened to the adjusting device (18, 18.2) are the two ends (29, 31) of a headband (28) which runs part of the way round on the inner side of the helmet opening (26),- the adjusting device (18. [CONT.]
	AT 324054	T			
	DE 50303117	D1			
	EP 1513424	A2			
	WO 2004000054	A2			
	WO 2004000054	A3			
	WO 2004000054	A8			

Patbase Publication Level data

- Several publications in a family may have interesting data elements, such as status events or sequences.
- Example: filter and summarize negative legal status events from publication level data

Family Legal Status:Neg	
Pub Number	Latest Legal Status
AU 200067713 A5	20020502: - (MK6) APPLICATION LAPSED SECTION 142(2)(F)/REG. 8.3(3) - PCT APPLIC. NOT ENTERING NATIONAL PHASE
CA 2190587 AA	20060519: - (FZDE) DEAD
EP 0760014 A1	20080813: - (18D) DEEMED TO BE WITHDRAWN (EFFECTIVE DATE : 20080123)
JP 10500575 T2	20060110: - (A02) DECISION OF REFUSAL (DESCRIPTION : JAPANESE INTERMEDIATE CODE: A02 EFFECTIVE DATE : 20060110)

Summarize GenomeQuest sequence results

- Report selected data for all hits in a family

	Title	Database	Patent Assignee	Query ID	Sequence Locations				
					Seq. ID Number	% Identity	Length	Location	
1.	PRODUCTION OF PEPTIDES IN PLANTS AS VIRAL COAT PROTEIN FUSION 1.1 Patbase	1.1 Patbase | link 1.2 GENESEQ | link	LARGE SCALE BIOLOGY CORP. 1.2 GENESE	query2	WO20050108564-0101	100.00	17	Example 6; SEQ ID NO 101; 115pp; English.	1.2
2.	Chimeric ebola virus envelopes and uses therefor 2.1 Patbase	2.1 Patbase | link 2.2 GPATPRT | link 2.3 GPATPRT | link 2.4 GPATPRT | link 2.5 GPATPRT | link 2.6 GENESEQ | link 2.7 GENESEQ | link	UNIV PENNSYLVANIA. 2.6 GENESE	query2 query3	US20050255123-0001	100.00	17	claim: 17	2.2
					WO03092582-0009	100.00	498	claim: 17	2.3
					WO03092582-0001	100.00	17	claim: 17	2.4
					US20050255123-0009	100.00	498	claim: 17	2.5
					WO20030092582-0001	100.00	17	Claim 17; SEQ ID NO 1; 107pp; English.	2.6
					WO20030092582-0009	100.00	498	Claim 17; SEQ ID NO 9; 107pp; English.	2.7
3.	ANTIGEN FRAGMENT AND TRUNCATION BASED ON EBOLA VIRUS ENVELOPE PROTEIN AS WELL AS APPLICATION 3.1 Patbase	3.1 Patbase | link 3.2 GENESEQ | link 3.3 GENESEQ | link	BIOENGINEERING RES INST ACAD MEDICAL SCI. 3.2 GENESE	query2	CN103864904-0008	100.00	17	Example 1; SEQ ID NO 8; 28pp; Chinese.	3.2
					CN103864904-0002	100.00	17	Example 1; SEQ ID NO 2; 28pp; Chinese.	3.3
4.	HUMAN EBOLA VIRUS SPECIES AND COMPOSITIONS AND METHODS THEREOF 4.1 Patbase	4.1 Patbase | link 4.2 GPATPRT | link 4.3 GPATPRT | link 4.4 GPATNUC | link 4.5 GPATNUC | link 4.6 GENESEQ | link	US DEPT HEALTH & HUMAN SERVICES. 4.6 GENESE	query7 query5	US20120251502-0011	100.00	9	claim: 8; 11; 12	4.2
					EP2350270-0011	100.00	9	TBD (information not in GQ-Pat)	4.3
					US20120251502-0027	100.00	20	probable disclosure (not found by automated parsing)	4.4
					EP2350270-0027	100.00	20	TBD (information not in GQ-Pat)	4.5
					WO20100048615-0027	100.00	20	Claim 30; SEQ ID NO 27; 98pp; English.	4.6

Integrating data over time

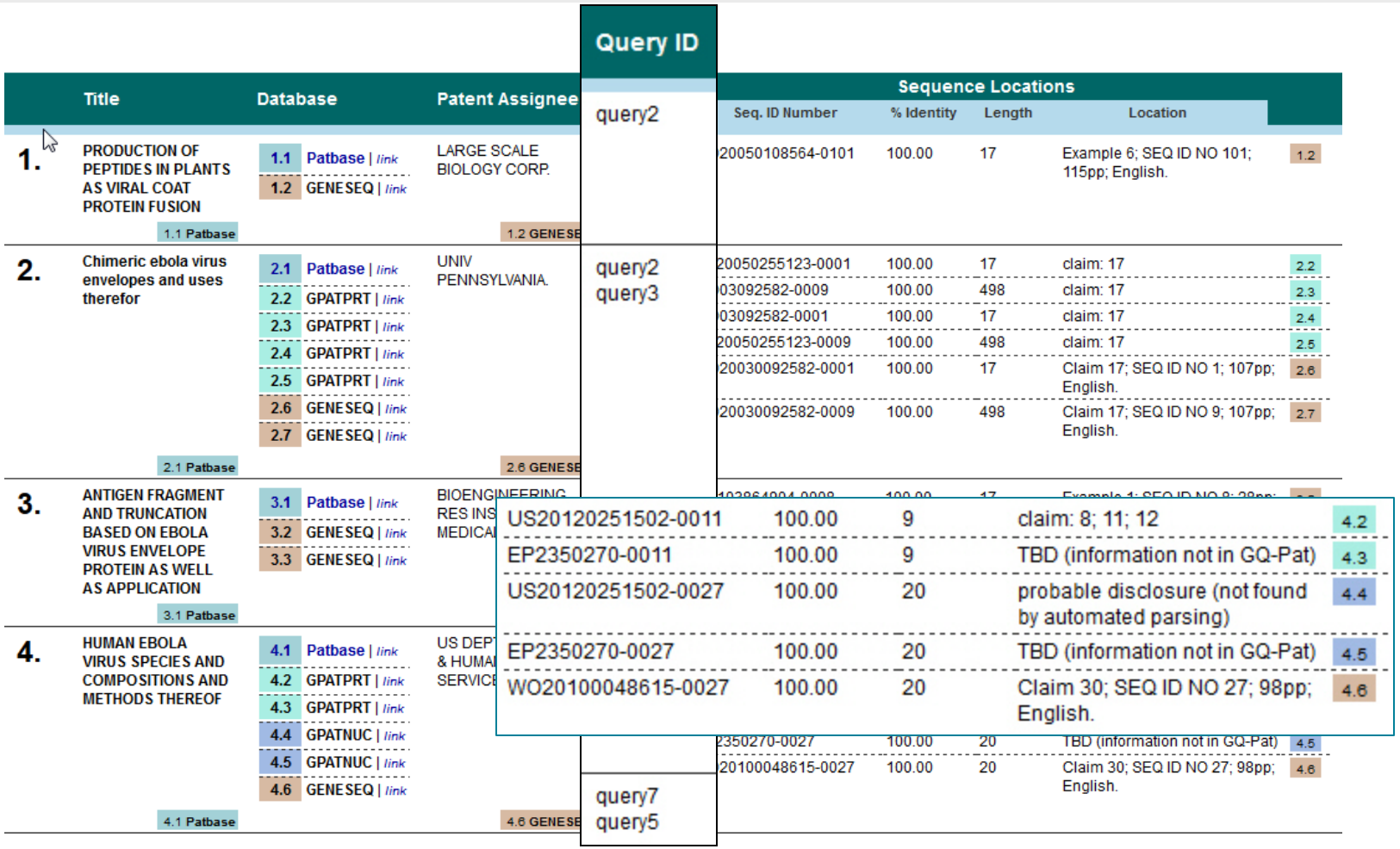
- Thomson Innovation has an option to return a single member for each family retrieved by a query.
- But the member returned may vary over time.
- Updates are grouped by family, sorted by most recent document in the family for analysis.

Integrate records from a single source

- Example: identify relevant changes in a search from Thomson Reuters Innovation

Database	Row Status	Pub. Number	Pub. Date	Patent Family		
				Patent	Kind	Date
Thomson Innovation + DWPI	Not in Added Results	WO 2013039855 A1	2013-03-21	US 20130064793	A1	2013-03-14
				WO 2013039855	A1	2013-03-21
				AU 2012308900	A1	2013-05-09
				CA 2847892	A1	2013-03-21
				EP 2755985	A1	2014-07-23
				CN 103906759	A	2014-07-02
				JP 2014526474	A	2014-10-06
Thomson Innovation + DWPI	Added	US 8951985 B2	2015-02-10	US 20130064793	A1	2013-03-14
				WO 2013039855	A1	2013-03-21
				AU 2012308900	A1	2013-05-09
				CA 2847892	A1	2013-03-21
				EP 2755985	A1	2014-07-23
				CN 103906759	A	2014-07-02
				JP 2014526474	A	2014-10-06
				US 8951985	B2	2015-02-10
Thomson Innovation + DWPI	Added	US 20130064793 A1	2013-03-14	US 20130064793	A1	2013-03-14
				WO 2013039855	A1	2013-03-21
				AU 2012308900	A1	2013-05-09
				CA 2847892	A1	2013-03-21
				EP 2755985	A1	2014-07-23
				CN 103906759	A	2014-07-02
				JP 2014526474	A	2014-10-06
				US 8951985	B2	2015-02-10

Revisiting summarized sequence results



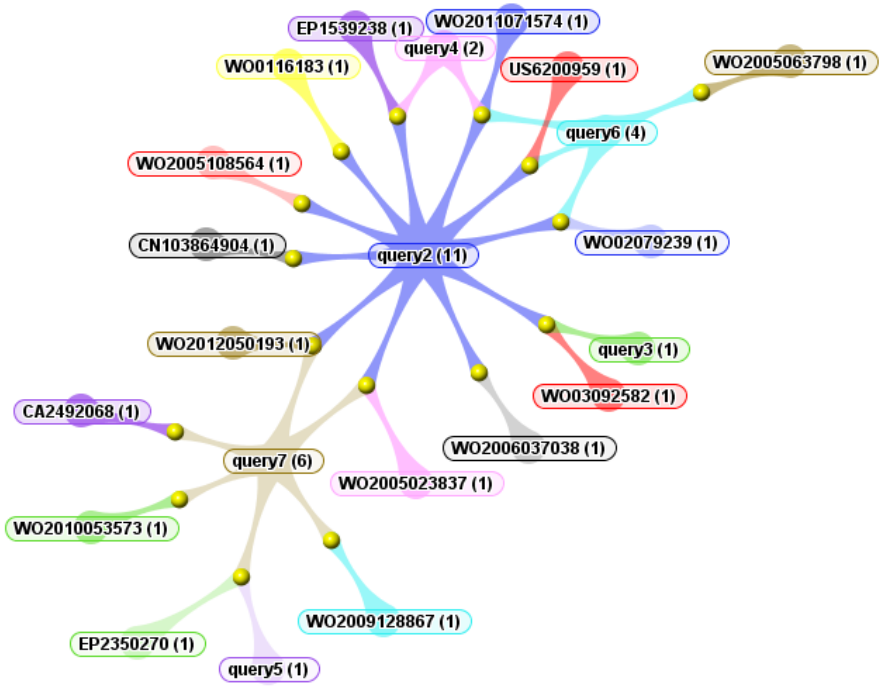
	Title	Database	Patent Assignee	Query ID	Sequence Locations				
					Seq. ID Number	% Identity	Length	Location	
1.	PRODUCTION OF PEPTIDES IN PLANTS AS VIRAL COAT PROTEIN FUSION	1.1 Patbase | link 1.2 GENESEQ | link	LARGE SCALE BIOLOGY CORP.	query2	20050108564-0101	100.00	17	Example 6; SEQ ID NO 101; 115pp; English.	1.2
2.	Chimeric ebola virus envelopes and uses therefor	2.1 Patbase | link 2.2 GPATPRT | link 2.3 GPATPRT | link 2.4 GPATPRT | link 2.5 GPATPRT | link 2.6 GENESEQ | link 2.7 GENESEQ | link	UNIV PENNSYLVANIA.	query2 query3	20050255123-0001 03092582-0009 03092582-0001 20050255123-0009 20030092582-0001 20030092582-0009	100.00 100.00 100.00 100.00 100.00 100.00	17 498 17 498 17 498	claim: 17 claim: 17 claim: 17 claim: 17 Claim 17; SEQ ID NO 1; 107pp; English. Claim 17; SEQ ID NO 9; 107pp; English.	2.2 2.3 2.4 2.5 2.6 2.7
3.	ANTIGEN FRAGMENT AND TRUNCATION BASED ON EBOLA VIRUS ENVELOPE PROTEIN AS WELL AS APPLICATION	3.1 Patbase | link 3.2 GENESEQ | link 3.3 GENESEQ | link	BIOENGINEERING RES INS MEDICA		US20120251502-0011 EP2350270-0011 US20120251502-0027 EP2350270-0027 WO20100048615-0027	100.00 100.00 100.00 100.00 100.00	9 9 20 20 20	claim: 8; 11; 12 TBD (information not in GQ-Pat) probable disclosure (not found by automated parsing) TBD (information not in GQ-Pat) Claim 30; SEQ ID NO 27; 98pp; English.	4.2 4.3 4.4 4.5 4.6
4.	HUMAN EBOLA VIRUS SPECIES AND COMPOSITIONS AND METHODS THEREOF	4.1 Patbase | link 4.2 GPATPRT | link 4.3 GPATPRT | link 4.4 GPATNUC | link 4.5 GPATNUC | link 4.6 GENESEQ | link	US DEPT & HUMA SERVIC	query7 query5					

Summarize sequence query locations

- Each sequence is tagged in the Query ID column (automatically imported from GQ)
- Or you can create a “Strategy” column and fill this column with the strategy name in the chart corresponding to that search.
- Combine results *without removing duplicates*.
- Use Summarize Unique Values rule in Reference Rows.

Integrate IP sequence results

- Unique dedup'd list of query hits allows us to visualize query results for each family



Searching multiple patent databases

- Searching multiple databases is a common strategy to increase recall, whether due to indexing, coverage, or timeliness.
- Whether results are presented from several sources, or answers are crossed to a single system, techniques to integrate family data are useful for producing better reports.

Software for
Business Intelligence

BizInt Smart Charts

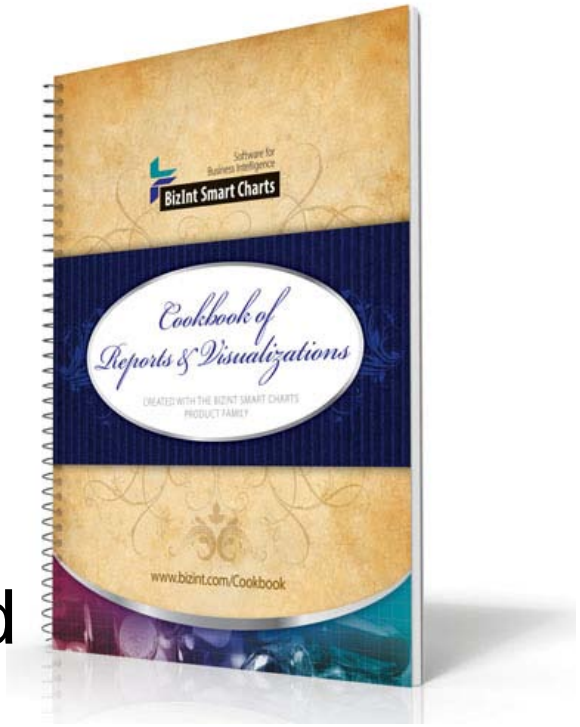
Thank you!

john.willmore@bizint.com

www.bizint.com

“Cookbook” of techniques

- Our **Cookbook** contains a collection of recipes to help you create accurate and appealing reports from integrated results.
- **Step-by-step recipes** illustrate how each tool can be used.
- Substitute ingredients as needed to answer *your* business questions.



BizInt Smart Charts
for Patents

VERSION 3.7

Patent Databases

Provide data on patents filed worldwide

- STN
- Questel Orbit.com
- Minesoft PatBase
- Thomson Innovation, Cortellis IP, Integrity Patents
- LexisNexis TotalPatent
- LifeQuest

BizInt Smart Charts
for Patents

VERSION 3.7

IP Sequence Databases

Provide data on sequences filed in patents.

- GenomeQuest (Geneseq, GQ-PAT)
- STN (USGENE, DGENE, PCTGEN)

Tools for integrating patent data

- **Combine charts** using File|Combine command
- **Identify related records** using the **"Identify Common Patent Family"** tool.
- Use **BizInt Smart Charts Reference Rows** to summarize related records in a single row.
- Clean-up and filter terms across records using **VantagePoint – Smart Charts Edition**.

BizInt Smart Charts
for Patents

BizInt Smart Charts
Reference Rows™

vantage point
Smart Charts Edition