Stuck between Quality and Cost Control: Challenge of the Patent Information Professional

Frederic Baudour

CFIB

Objectives of the Presentation

PIP

(Patent Information Specialists)

Freedom to Operate

Patentability

Invalidation

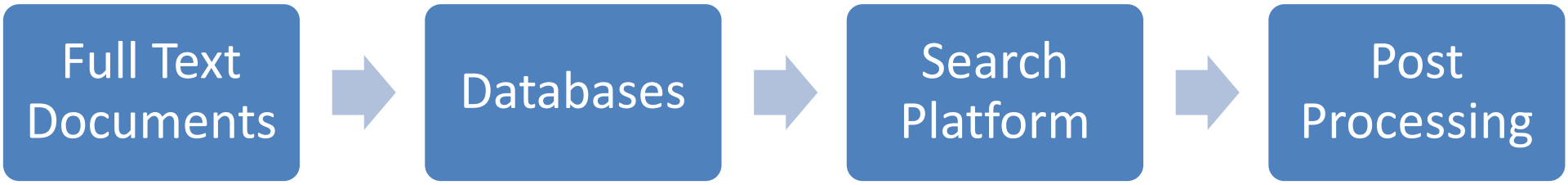
Patent Landscape

...

Need to deliver Quality

Budget

Quality is directly related to …
Information Value Chain



Patents
Non-Patents

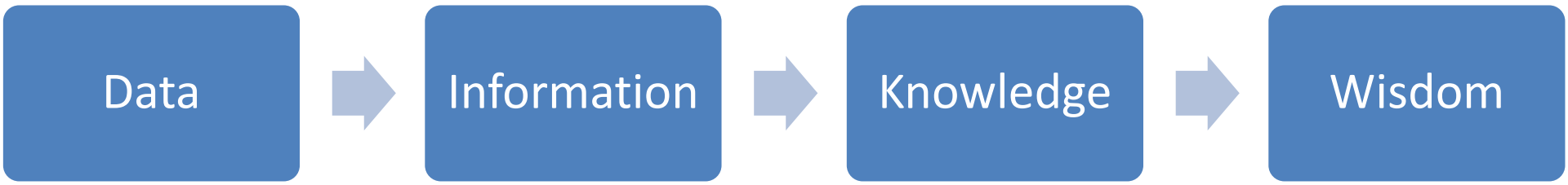
Added value
Full Text

Functionalities

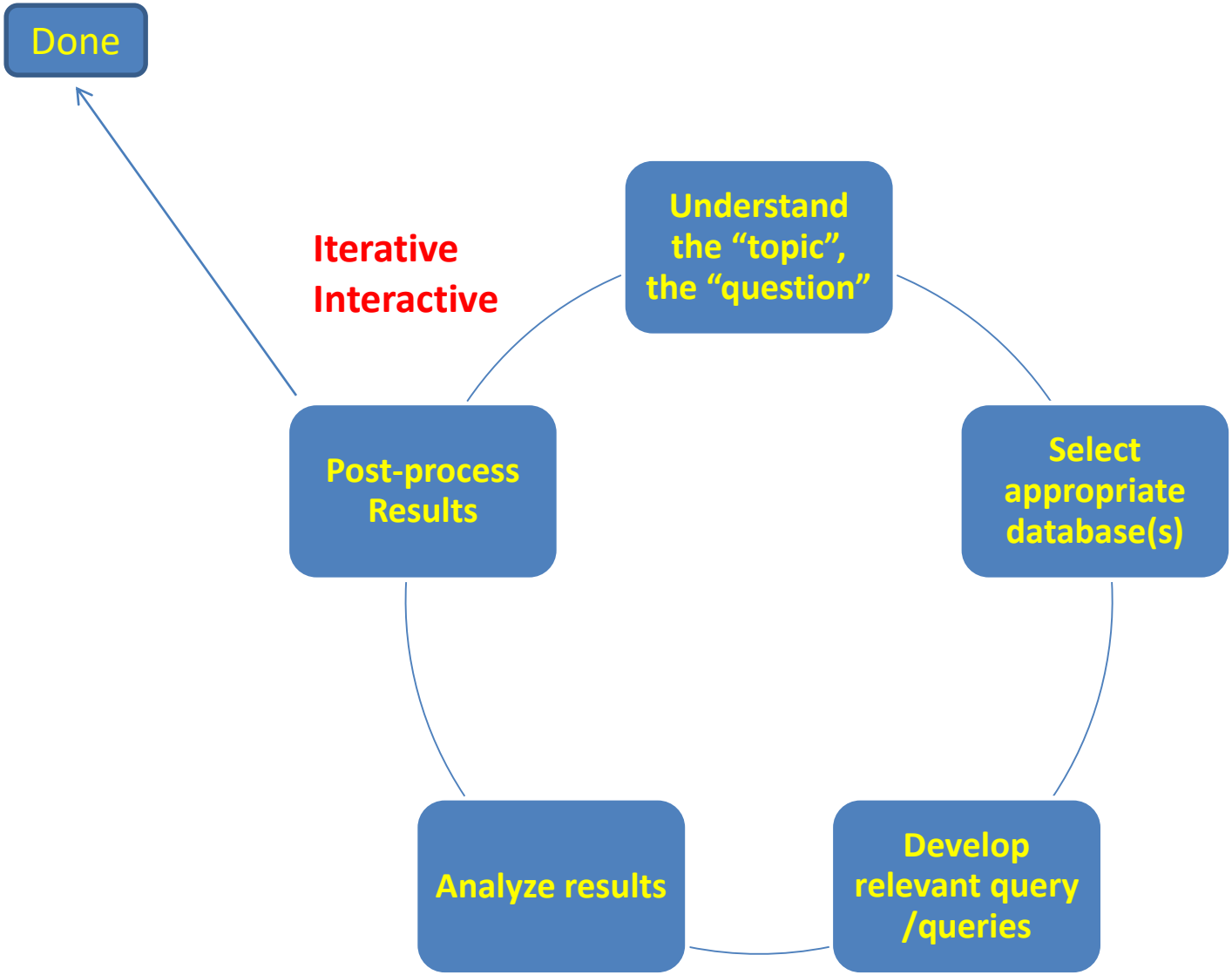
Selection of Documents

Analysis & Visualization

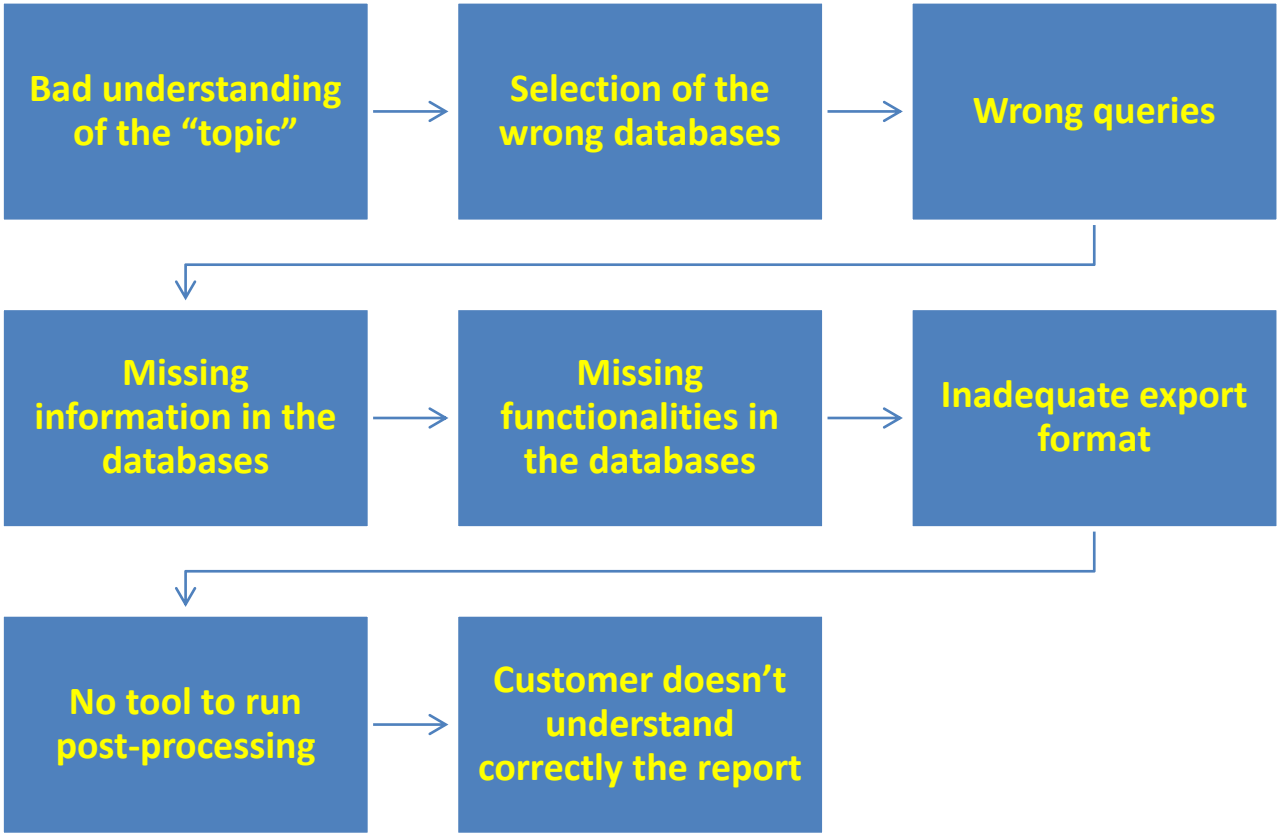
Others



PIP: Usual Work Process



Where can it go wrong?



But there is another parameter !!!

Cost of Knowledge vs Cost of not Knowing

Everything has a cost !

You cannot subscribe to all the databases available in the World

Free databases, Full texts or added value ones ???

You cannot buy all the Analysis & Visualization tools available in the World

But each one has some specific features

Your time has a cost

And your Management never speaks about budget increases !

What is your service Return of Investment?

Everything has a cost !

Missing a relevant patent / document has a high cost too !!!

How should we select our “tools”?

Selecting the tools you will use depends:

- On your needs,
  - On their strengths,
  - On their weaknesses,
- On their price!

Need to have

vs

Nice to have

Investment: You don’t subscribe just for 5 min

Need also to consider the tool integration in your environment, training, change mgt, …

The Money "Value" Chain

	I have no money	I have a quite limited budget	I have a little bit more money	I am rich
Where?	Free databases: -Espacenet -WIPO -USPTO -…	"Full Text" Databases: -PatBase -Orbit -TotalPatent -PatSeer -…	A mix of full text databases and added value ones (Derwent, STN, …)	Who cares? You are rich !
Advantages	You can already find a lot of documents for free	You can search full text documents, with patent family concept	You can use the best of both worlds	You are rich!
Defaults	Need to repeat the search in different databases Limited patent family concepts Never fully comprehensive	You don't have access to some added value indexing and/or functionalities You sometime need to be creative	Need to juggle between added value db's and full text ones in order to keep costs under control	None: you are rich!

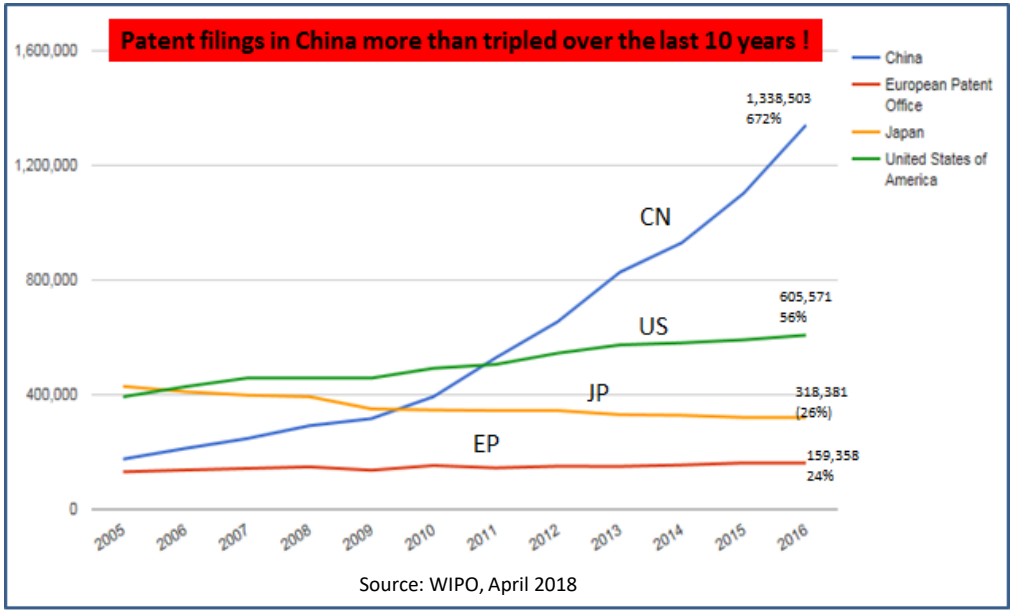
Despite how much money you have, you always need to work carefully !

Understanding what are the strengths and weaknesses of the Information Value Chain will help you to select appropriate tool(s) according to YOUR needs

Problems: Original Documents

- Different languages
- Different writing styles
- Different technology domains
- All the information is not always available (eg patent republication in some countries)
- “Explosion” of the number of Chinese Patents !!!
- Volatile information, eg Internet page citations

Consequences on <u>WHERE</u> you will be searching and <u>HOW</u> you will be searching



Source: WIPO, April 2018

Problems: Databases

- "Full Text Databases"
  - Original titles & abstracts
  - Patent Machine Translation
  - No added indexing
  - Non-Patent Literature (NPL)
- "Added value databases"
  - No full texts
  - Non-Patent Literature (NPL)
    - Per technical domain
  - Translations
  - Costs !
- Both:
  - Comprehensive search for all the patents from company X:
    - Added value db's < Full Text db's < INPADOC ???
  - Patents: some countries still not covered by databases

CN108205645A Machine translation: The present invention relates to a heterologous image matching system reference image quality evaluation method, for the preparation of the reference chart of satellite remote sensing image analysis and image feature matching to a heterologous image matching system to match the impact that the impact performance of the established benchmark figure characteristic elements; Design Reference figure spatial variation model, to accurately calculate the baseline figure features in geometry of observation of deformation; Gradient direction histogram description reference figure, reference image build multi-dimensional feature vector; Design Reference Figure multilevel encoded using statistical image sample set image matching reference image quality evaluation system performance, establish the calibration library; Using the similarity measure mutual information method to calculate the degree of similarity between the reference map, to find the most similar reference image reference image quality calibration library, to obtain the corresponding recognition performance matching system calibration values. The present invention ensures that in matching system work in unknown environment before, using the reference Fig. calibration is being prepared reference image library objective evaluation, guarantee high quality output reference Fig.

Problems: Functionalities

Added Value Functionalities

- Assumption that functionalities depend on the nature & quality of the data indexation
  - Eg Inventors, assignees standardization
  - Eg Patent transfers, change of assignees, …
  - Eg patent classifications
    - Which already depends on the indexer !!! (or the machine behind?)
  - Eg Patent Family definition are different in Derwent, Orbit, Inpadoc, …
  - Eg CAS numbers, chemical structures
  - Searching all cited patents for a specific patent collection → possible but then, who is citing what?

Basic Functionalities

- Issues with some browsers (IE, Chrome, Firefox)
- Document exports (eg MS Word Exports)
- …

Analysis & Visualization Tools

	Basic	Advanced
Generic	Lot of basic analysis rapidly available Some level of flexibility but limited Still "Black Boxes" Sometimes included in tool subscription eg: PatBase, Orbit, …	Lot of flexibility in analysis Your creativity is the limit Need to fully understand the tool you are using - Need time Software cost, database cost eg: VantagePoint (⇔ more than patents)
Specific		Using the full added value of a specific database eg: Anavist (costs based on the data volume)

New analysis concepts to mix different kinds of data: eg patent data with financial ones

Precise financial information?

Many parameters (eg market size, other players, competitive advantages, emerging tech, …

"Free" "open source" resources

- OpenRefine http://openrefine.org/
- Gephi https://gephi.org/

100% Flexibility
Can analyze very deeply
No Analysis/visualization cost

Need Expertise
Need Creativity
Need Time

Problems: Analysis & Visualization Tools

Quality of analysis depends on quality of the data !!!

Statistical analysis without qualitative analysis:
How do you evaluate the quality of your recordset???
Rubbish in => Rubbish out

Last few years:

Lot of new analysis & visualization concepts & tools but (a) basic improvements with data quality still needed and (b) added value still to be proven

From several survey with different user groups, MS Excel remains a tool used by (nearly) everybody when it comes to Analysis & Visualization

In-house post-processing

Be creative, You can still do a lot more!

Example: Who is citing the patents from company x?

In PatBase:

pa=x => y patent families (query 1)

Ctf 1 => z patent families (query 2)

Looking in a specific patent family from query 2: which patent(s) from query 1 is it citing?

Not clear from PatBase but possible to post-process in Excel

This patent	with title	from company	cites ths patent	with title
US20110391 11A	CURABLE FIBERGLASS BINDER ;	JOHNS MANVILLE	US2010197856A	SELF-CROSSLINKING BINDERS ;
US20102987 38A	VESSEL, COATING, INSPECTION AND PROCESSING APPARATUS ;	CV HOLDINGS LLC ;	US6087469A	POLYESTER POLYOLS OF LOW MOLAR MASS, THEIR PREPARATION AND USE IN COATING COMPOSITIONS ;
US20142750 48A	PRODRUGS OF FUMARATES AND THEIR USE IN TREATING VARIOUS DISEASES ;	ALKERMES PHARMA IRELAND LTD	US6166220A	Ethylenically unsaturated imidazidolidinone monomers ;
US20092989 62A	PHOTOLATENT BASES FOR SYSTEMS BASED ON BLOCKED ISOCYANATES ;	CIBA HOLDING INC ; CIBA SC HOLDING AG ; CIBALTD ; BASF SE	US2006051591A	RADIATION-CURING BINDERS AND A PROCESS FOR THEIR PREPARATION ;

Conclusions

- Be careful: There are still unsolved issues in the whole information value chain
- There are no “best” tool, only the one best suited for you:
  - Need to have ⇔ Nice to have
  - Control costs
- Important role of Patent Information Professional:
  - Nothing good from “automatic”
  - Balance & Control
  - Innovate: add value to your work !