Patinformatics: Tasks to Tools

Anthony Trippe
Science IP / CAS

What is Patinformatics?

- Patinformatics
  - the science of analyzing patent information to discover relationships and trends, which would be difficult to see when working with patent documents on a one-on-one basis. The term encompasses all forms of analyzing patent information including:

- Inventing Words
  - "I will concoct new phrases and definitions so that my ideas are revolutionary. As you know, nothing worth knowing can be explained with regular words." - Scott Adams, Dilbert and the Way of the Weasel

What is Patinformatics?

- Patent Intelligence - the use of patent information to identify the technical capabilities of an organization and the use of that intelligence to develop a strategy for strategic technical planning
- Patent Mapping - sometimes described as white space mapping, which uses published patent data to create a graphical or physical representation of the relevant art pertaining to a particular subject area or novel invention

What is Patinformatics?

- Patent Citation Analysis - the study of patent citations for potentially determining a patent's value or, perhaps more reliably, the identification of potential licensing partners or leads based on the citation of an organization's patents by another company in the same or a completely different market space
- Patinformatics can also cover additional applications of patent information involving a subsequent analysis step. The key underlying property in each of these diverse areas is the analysis step.

Patinformatics Philosophy

Law of Linear Patent Analysis

- Create a tool kit of patinformtics tools
- Understand the Need Behind the Need
- The Need Drives the Question
- The Question Drives the Data
- The Data Drives the Tool
- Why is this important?
  - To a man with a hammer, everything looks like a nail - avoid this at all costs

Tasks to Tools

- Full article appears in World Patent Information 25 (2003) 211-221

Tasks to Tools

- Technique
  - List Cleanup & Grouping of Concepts
- Definition
  - Manual or automatic standardization of terms within a data field. List cleanup is required in order to produce statistically relevant results.
- Utility
  - Grouping allows synonymous terms to be combined together so that their true value in a data set can be accurately assessed.
- Tools
  - VantagePoint, MS Excel, ClearResearch, OmniViz, Aureka ThemeScape

Tasks to Tools

- Technique
  - List Generation (Histograms)
- Definition
  - Provides counts of various patent related metrics within individual data fields.
- Utility
  - Allows the statistical comparison of two or more entities in the same data field.
- Tools
  - VantagePoint, MS Excel, ClearResearch, Aureka Reporting, Knowledgist, Technology Watch, Wisdomain Analysis Module, Delphion PatLab II, SciFinder

Tasks to Tools

- Technique
  - Co-Occurrency Matrices & Circle Graphs
- Definition
  - Data fields are placed on an X and Y-axis or on opposites sides of a circle. Number of overlapping occurrences of shared X and Y can be seen as numbers within the matrix or as lines of increasing width connecting items on the circle.
- Utility
  - Allows connections to be made between two or more fields of information and provides a representation of how strong the connection is.
- Tools
  - SciFinder Panorama, VantagePoint, ClearResearch, Aureka Reporting, Wisdomain Analysis Module, Delphion PatLab II

Tasks to Tools

- Technique
  - Clustering of Structured (Fielded) Data
- Definition
  - Intellectually assigned classification systems produce a standardized code that can be used as a means of organizing documents that share a similar coding structure.
- Utility
  - Documents, which share a high percentage of codes in commons, are likely to be similar. Allows a large number of documents to be organized.
- Tools
  - Technology Watch, ClearResearch, OmniViz, VantagePoint

Tasks to Tools

- Technique
  - Clustering of Unstructured (Text) Data
- Definition
  - Raw text is processed to identify concepts and phrases contained within. As with the clustering of structured data, concepts instead of codes are used to group documents that share a high degree of overlap.
- Utility
  - Documents, which share a high percentage of concepts in commons, are likely to be similar. Allows a large number of documents to be organized.
- Tools
  - Aureka ThemeScape, ClearResearch, OmniViz, Vivisimo, Delphion Text Clustering, VantagePoint

Tasks to Tools

- Technique
  - Mapping Document Clusters
- Definition
  - Document clusters are arranged in 2-dimensional space creating a map. Collections of documents, which share elements in common, are placed closer together geographically while collections with less similarity are placed further away.
- Utility
  - Allows relationships between clusters to be identified. Creates a visual representation of a document collection from a high-level view.
- Tools
  - Aureka ThemeScape, Technology Watch, ClearResearch, OmniViz, VantagePoint

Tasks to Tools

- Technique
  - Adding Temporal Component to Cluster Map
- Definition
  - A time dimension can be called out on a map usually by means of alternate colors.
- Utility
  - User can follow the progression of a subject as it develops or evolves.
- Tools
  - Aureka ThemeScape, ClearResearch, OmniViz

Tasks to Tools

- Technique
  - Citation Analysis
- Definition
  - When patent documents are examined relevant prior art is mentioned on the search report or on the front page of the documents. The number of citations can be counted or followed as they link documents together.
- Utility
  - Hyperbolic trees are used to show relationships between patents that cite one another. Citation counts are used to discover potentially pivotal documents.
- Tools
  - M-CAM, Aureka Citation Trees, Delphion Citation Link, Metrics Group Citation Bridge, Wisdomain Citation Module

Tasks to Tools

- Technique
  - Subject/Action/Object (SAO) Functions
- Definition
  - Parts of language that are used to describe the teachings that the author wants to portray. Key SAOs encapsulate the technical learnings contained in a document.SAOs can be described as problems and solutions.
- Utility
  - By identifying SAOs the teachings of a document can be isolated and examined from the rest of the document creating a knowledge base
- Tools
  - Knowledgist

Tasks to Tools Matrix

Tool	List Cleanup & Grouping of Concepts	List Generation	Co-Occurrency Matrices & Circle Graphs	Clustering of Structured (Fielded) Data	Clustering of Unstructured (Text) Data	Mapping Document Clusters	Adding Temporal Component to Cluster Map	Citation Analysis	Subject/Action/Object (SAO) Functions
Aureka ThemeScape	X				X	X	X		
ClearResearch	X	X	X	X	X	X	X		
OmniViz	X			X	X	X	X		
Vivisimo					X				
Delphion Text Clustering					X				
Technology Watch		X		X		X			
M-CAM								X	
VantagePoint	X	X	X	X	X	X			
Aureka Citation Trees								X	
Delphion Citation Link								X	

Task to Tools Matrix

Tool	List Cleanup & Grouping of Concepts	List Generation	Co-Occurrency Matrices & Circle Graphs	Clustering of Structured (Fielded) Data	Clustering of Unstructured (Text) Data	Mapping Document Clusters	Adding Temporal Component to Cluster Map	Citation Analysis	Subject/Action/Object (SAO) Functions
Metrics Group Citation Bridge								X	
Wisdomain Citation Module								X	
Knowledgist									X
MS Excel	X	X							
Aureka Reporting		X	X						
Wisdomain Analysis Module		X	X						
Delphion PatLab II		X	X						
SciFinder		X							
SciFinder Panorama			X						

- Analysis
  - "The word analysis is formed by the root word anal and the ancient Greek word ysis, meaning "to pull numbers from"". - Scott Adams, Dilbert and the Way of the Weasel

- Forecasting
  - "When you lie about the future, that's called optimism, and it is considered a virtue. Technically speaking you can't "lie" about the future because no one knows what will happen. When you apply this unique brand of optimism (not lying!) at work, that's called forecasting." - Scott Adams, Dilbert and the Way of the Weasel

One Last Thing

Our First Continuation-in-Part (CIP)