Mapping the innovation terrain: exploring patent analytics and the evolving role of the analyst

#### **Chris Harrison**

Patent Analytics Manager, WIPO

EPO Patent Knowledge Week November 29, 2023



#### Chris Harrison

- MEng Systems Engineering, Cardiff University, UK
- UK Intellectual Property Office 2008-2023
   Patent Analyst since 2010
   Head of IP Analytics and Data Insights (2017-2023)
- Joined WIPO as Patent Analytics Manager in 2023
- Qualified Patent Information Professional (QPIP)
   13+ years' experience in patent information
- <u>christopher.harrison@wipo.int</u><u>linkedin.com/in/chris-harrison-wipo</u>



## What is patent analytics?

Leveraging the wealth of knowledge contained in patent data to provide added-value information and supporting informed decision-making

## What is patent analytics?

Analyzing different patent fields to identify technology trends, key and emerging players, existing and potential markets, opportunities and collaboration

Scope varies – from very narrow to very broad

Type of analysis and insights vary depending on the specific needs and questions to be answered

- Patent Landscaping
- Patent Analytics
- > Patent Knowledge
  - IP Analytics
  - > IP Intelligence



# Patent Analytics Market Size to Surpass USD 2,364.6 Million by 2030, exhibiting a CAGR of 13%

As per the report by Fortune Business Insights, the global Patent Analytics Market size is projected to reach USD 2,364.6 Million by 2030, at a CAGR of 13% during the forecast period.

May 24, 2023 09:00 ET | Source: Fortune Business Insights

Follow

## Target audience



Government



Business



**Investors** 



Researchers

## Patent analytics for government

Do we have start-ups to support?

# How to develop our industrial policy?

What are the technologies to come?



# Patent analytics for business including start-ups

What partnership opportunities are available?

# What new technology should be patented?

What are the strengths and weaknesses of our own patent portfolio?

Where should we invest our R&D resources?



## Patent analytics for investors

Which companies are most active in patenting of a particular technology?

# What are current technology trends?

Who is actively acquiring patents in the sector?



## Patent analytics for inventors, researchers and universities

What new technology should be patented?

What is the leading technological solution in the area?

with to

What companies can we partner with to commercialize our research?

# Unlocking the value of patent data

Patent analytics presents patent data, search results and findings in a **structured way** 

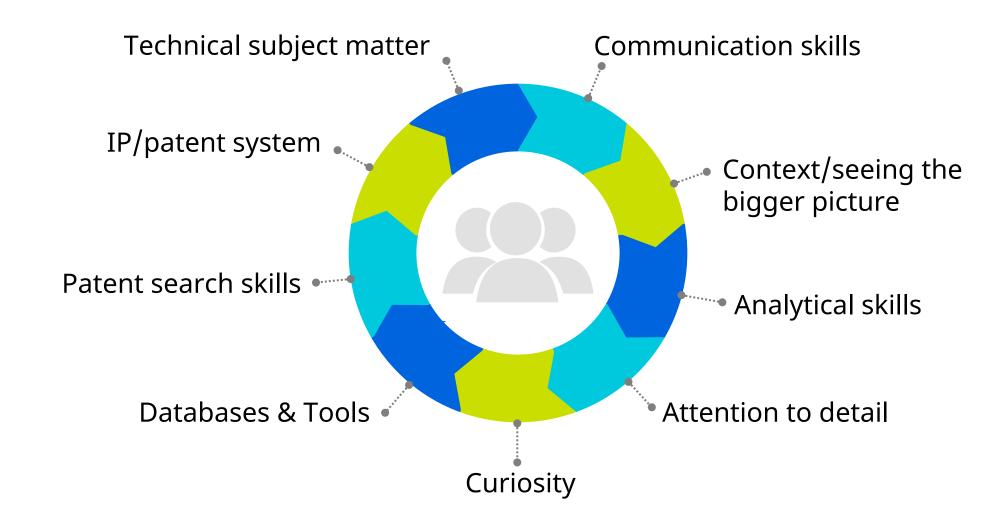
Creates an output that makes technical data easily understood by non-experts

Includes visualization of the results making the information easy to digest

Facilitates interdisciplinary dialogue between various stakeholders



## The key skills of a patent analyst



## Patent database and analytics tools providers

#### Free-of-charge patent databases **PATENTSCOPE** 101 **Espacenet** Patent search

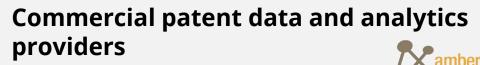




R Studio

















Patent iNSIGHT Pro









#### **Commercial analytical tools**













13

## Open source tools

Typically used by researchers, data scientists, IT and business analysts

#### **Pros:**

- Free of charge
- Customizable to specific needs
- Give flexibility to experiment and choose the tool(s) that is more adapted to the user needs
- Support / troubleshooting by programmers and other users in various online fora

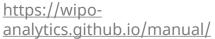
#### Cons:

- Not as intuitive or userfriendly as commercial tools
- Require some programming skills, time and practice
- Several bugs/problems
- No commercial support

Keep voting in the live poll! Slido.com #2836777









## The patent analytics process



Stage 2 Patent search Stage 3 Data cleaning/ normalization Stage 4
Data analysis
and visualization

Stage 5 Narrative and storytelling

Stage 6
Dissemination

## Topic and scope definition

#### The importance of a clear project specification

- Detailed understanding of the business need (the question behind the question)
- Brainstorming and discussion to better define the specification
- Sufficient background information
- Defining project scope and the most useful and appropriate analysis an iterative process

#### **Managing expectations**

- Options and limitations of patent analytics
- Timelines and feasibility

Stage 1
Topic and
scope definition

Stage 2 Patent search Stage 3 Data cleaning/ normalization Stage 4 Data analysis and visualization Stage 5 Narrative and storytelling

Stage 6
Dissemination

#### Patent search

#### Selecting the most appropriate data source

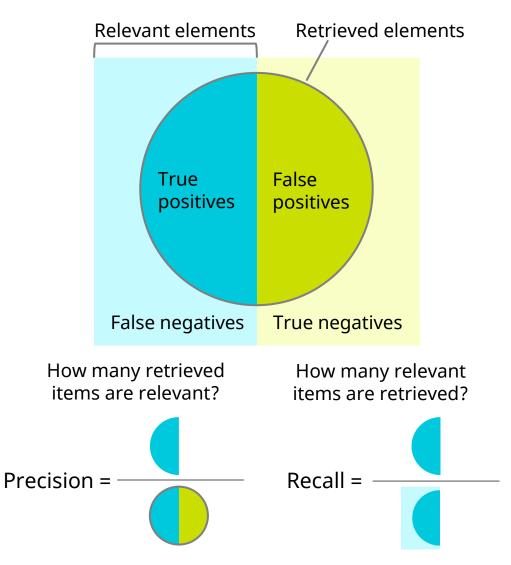
Required geographical and historical coverage

#### **Iterative search strategy**

- Preliminary searches to explore the field and define approach (classification and/or keyword)
- Search strategies need to be tailored and refined such as to optimize precision and recall

#### Minimize dataset noise

 Aim is to eliminate false positives and false negatives (although the reality is the minimize these)



Stage 1
Topic and scope
definition

Stage 2 Patent search Stage 3
Data cleaning/
normalization

Stage 4 Data analysis and visualization Stage 5 Narrative and storytelling

Stage 6
Dissemination

## Data cleaning / normalization

#### A critical step

- How does the exported data look? Applicant / inventor name variations and misspellings, patent number normalization etc
- Clustering / grouping (by affiliation, M&A information etc)
- Manual or semi-automatic (commercial data providers often have their own data cleaning tools built-in that aim to reduce the data cleaning workload)
- If you do not clean your data then any analysis derived from it will be wrong (applicants, inventors, CPC/IPC)



Stage 2 Patent search Stage 3 Data cleaning/ normalization Stage 4 Data analysis and visualization Stage 5 Narrative and storytelling

Stage 6
Dissemination

## Data analysis and visualization

#### Allow yourself to be led by the data

- Avoid using a predefined list of charts/graphs
- Make sure the analysis covers the questions to be answered

#### **Technology monitoring**

- ✓ Which technology trends exist and how have they developed over time?
- ✓ Where are the crowded areas?

#### **R&D** decision-making

- ✓ Are there any gaps or white spaces, i.e. areas with little patent protection, that permit business opportunities?
- ✓ What further applications or uses are possible?
- ✓ Which further adaptations or embodiments could be explored?
- ✓ Identify solutions for technical problems

#### **Competitors monitoring**

- ✓ Which players are the most active?
- ✓ Patent portfolio of competitors?
- ✓ Who bought or sold IP rights?

#### **Investment and collaboration decisions**

- ✓ Which other patents are most relevant for our own activities?
- ✓ Check for potential infringements of IP rights (FTO, product clearance; in licensing)
- ✓ Monitor expiry of protected technologies for later use
- ✓ Identify business opportunities (out-licensing; potential partners with know-how; marketing opportunities)

Stage 3
Data cleaning/
normalization

Stage 4
Data analysis
and visualization

Stage 5 Narrative and storytelling

Stage 6
Dissemination

## Data analysis and visualization

#### Simplicity is a virtue

- Humans naturally focus on information that stands out (peaks etc), and we tend to look for patterns and draw our own conclusions
- The same information can be presented in multiple ways so experiment to see which is the most appropriate, and more self-explanatory visualization

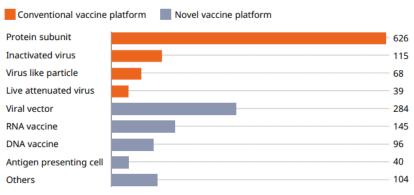


# "Everything should be made as simple as possible, but no simpler"

Albert Einstein, 1879

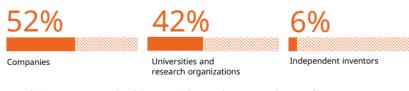
#### Distribution of patent applications across vaccine Platforms

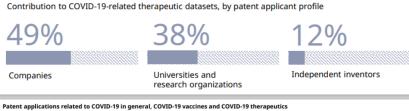
48% of the vaccine patent dataset related to protein subunit vaccines, followed by patent filings related to viral vector vaccines), which accounted for 22% of the vaccine patent dataset.

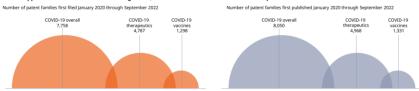


## Patenting activity related to COVID-19 vaccines and therapeutics is nearly equally distributed between companies and universities/research organizations

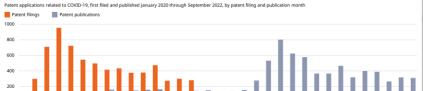
Contribution to COVID-19-related vaccine datasets, by patent applicant profile







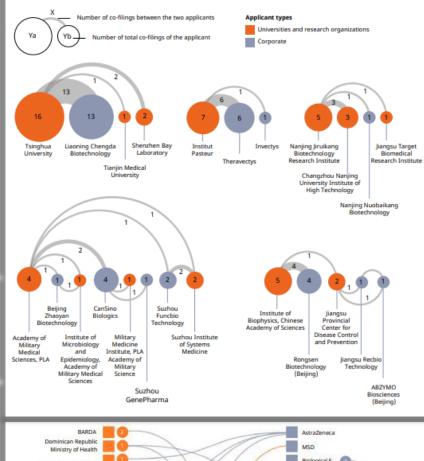
#### COVID-19-related patent application filing peaked in April 2020, followed by a corresponding peak in COVID-19-related publications in October 2021





**Figure 11.** Network analysis of the top collaborating entities in COVID-19 vaccine patent applications

Among COVID-19 vaccine patents, Chinese companies and research institutions show greatest cooperation, followed by French corporates and research institutions.

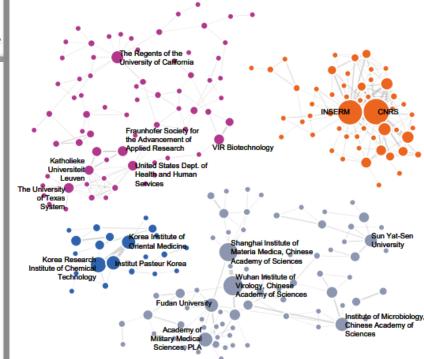


BioNTech
CureVac (1) (3)

European Commission

Figure 15. Network analysis of top collaborating entities in COVID-19 therapeutic patent applications

Analysis shows four main collaborating communities: a community primarily composed of United States and European universities and research institutions (purple), a group centered around two French research institutions, CNRS and INSERM (orange), a group centered around research institutions from the Republic of Korea (blue) and a group primarily composed of Chinese corporate entities and research institutions (gray).



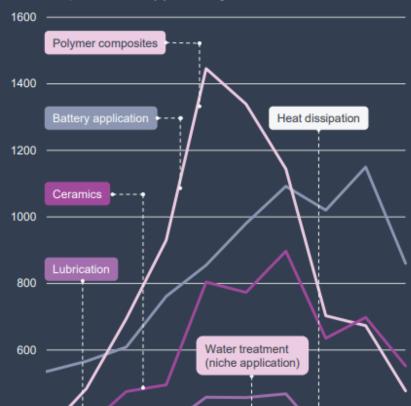


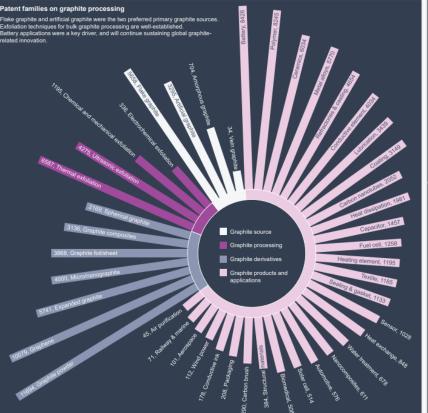


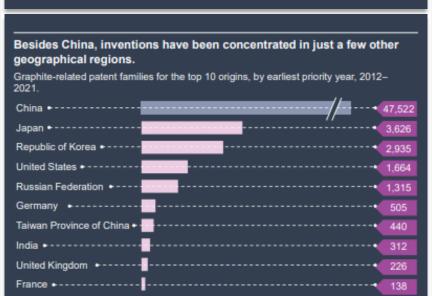
China dominated patent filings, with more than 47,000 patent families



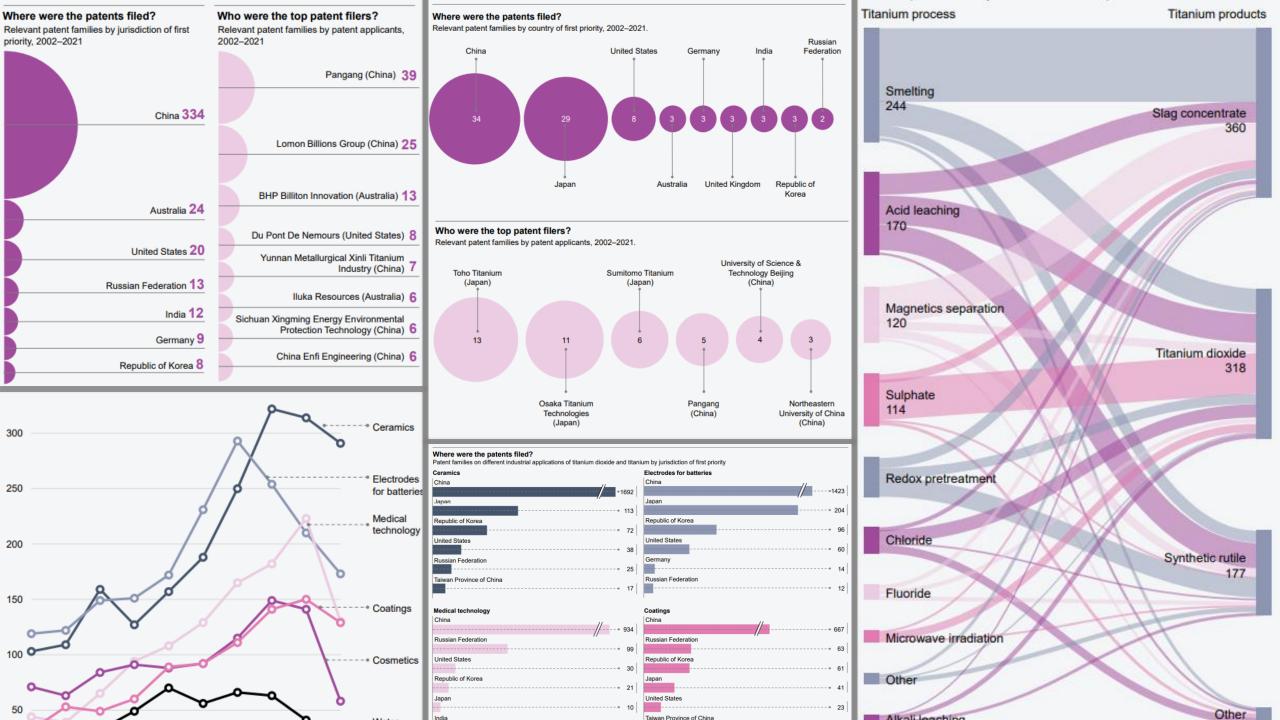
Number of patent families by year of filing, 2012–2021.



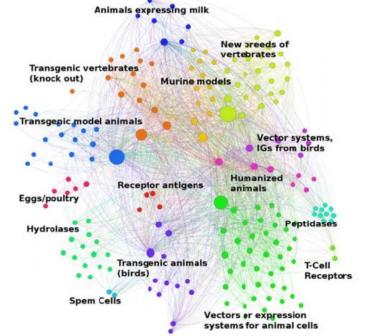


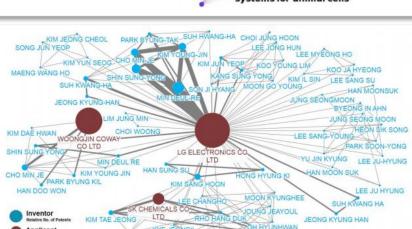


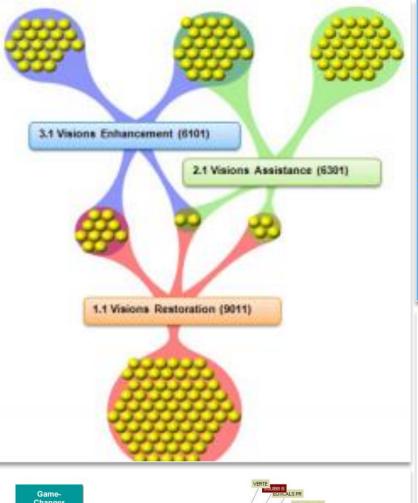


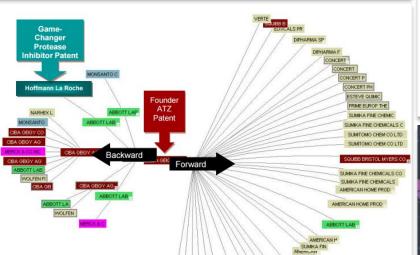


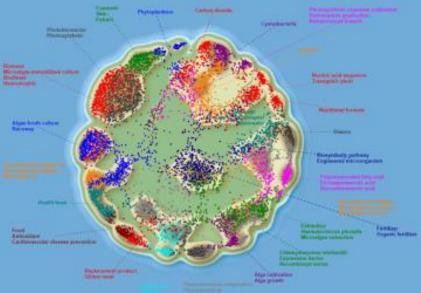




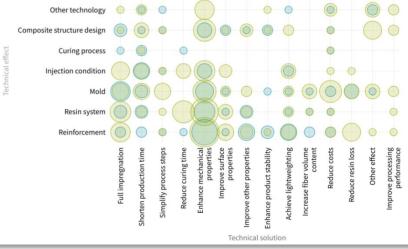














Stage 4 Data analysis and visualization Stage 5 Narrative and storytelling

Stage 6
Dissemination

## Narrative and storytelling

#### Tailored to the audience

- Report style, language and length should be adapted to the target audience
- Multiple reader profiles (e.g. lawyers, managers, scientists)
- The report should be easy to read, the messages and results clear and the visualizations intuitive and meaningful

#### Bringing the report to life

- An accompanying infographic, a storytelling approach, and/or a dashboard may be appropriate based on specific objective of the report
- Do not underestimate the importance of this step to ensure your report lands well

Stage 5 Narrative and storytelling

## Narrative and storytelling



















Data

27

Sorted

Arranged

Presented visually

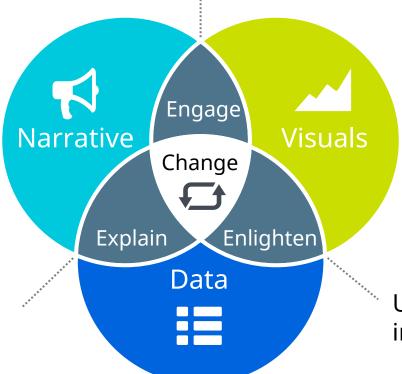
Explained with a story

Stage 4 Data analysis and visualization Stage 5 Narrative and storytelling

Stage 6
Dissemination

Narrative and storytelling

Combine narrative with visuals to connect with your audience.



Help your audience interpret and understand your insights.

Use data visualizations to reveal insights hidden in the data.

Stage 3
Data cleaning/
normalization

Stage 4
Data analysis
and visualization

Stage 5 Narrative and storytelling

Stage 6
Dissemination

## Dissemination

A valued and useful resource, not a dust collector



#### **Content Sharing**

- Accessible reporting
- Presentation vs. sharing only a written report
- Static vs. interactive
- Infographics
- Background data / dataset / methodology



#### **Target Audience**

- Publicly available vs. shared privately
- Sharing with end-users of the report or intermediaries



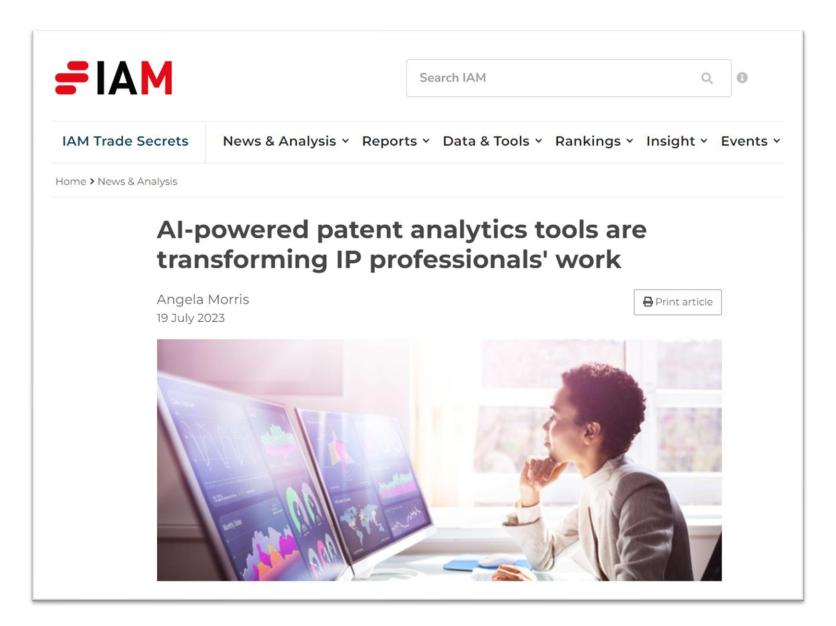
#### **Distribution Channels**

- Social media
- Website
- Conference / seminar / workshop

# The rise of AI-powered patent analytics tools

- The number of patent analytics platforms utilizing artificial intelligence is proliferating
- AI algorithms that run them have developed substantially in recent years
- Vast quantities of patent data can now be analyzed much more quickly



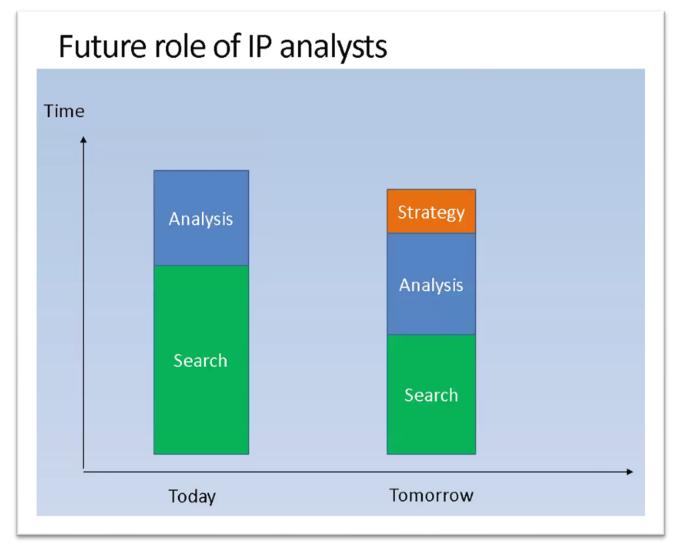


# AI tools are changing the patent analyst's role

- AI tools will not replace the analyst
- But the use of AI tools will allow day-to-day tasks to be completed much faster
- And potentially deeper insights will be uncovered to better inform decision making
- AI tools are already very effective at speeding up data cleaning
- Massive improvements made in AI-assisted patent search too
- ChatGPT data storytelling?



## Improving resource efficiency

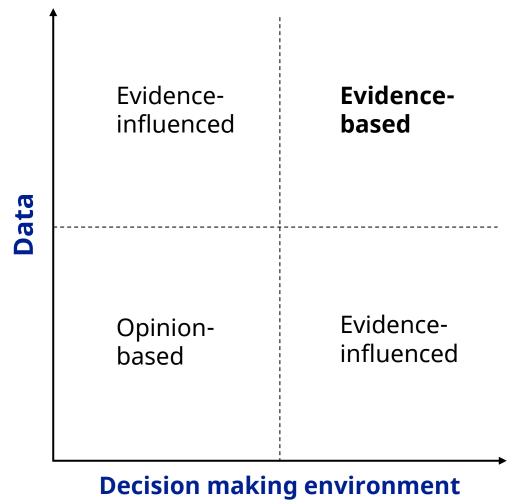




Yuval Noah Harari, 2018

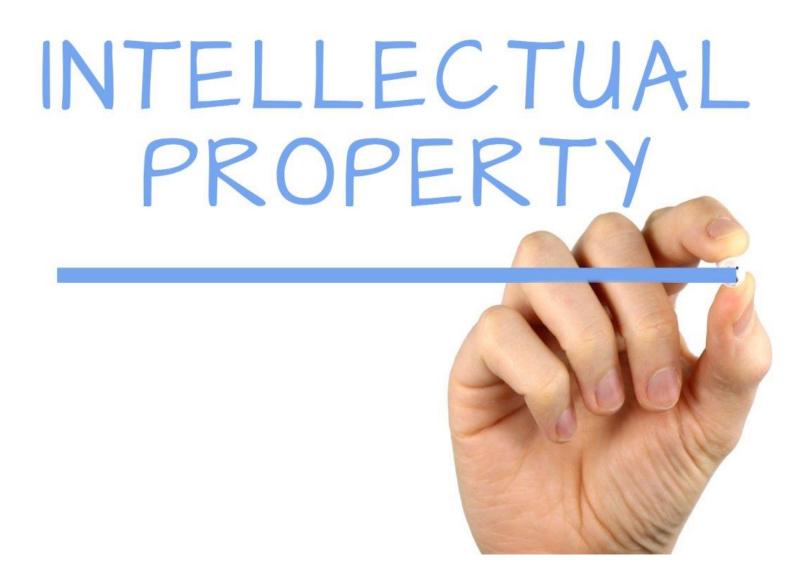


Why evidence-based decision making is important?

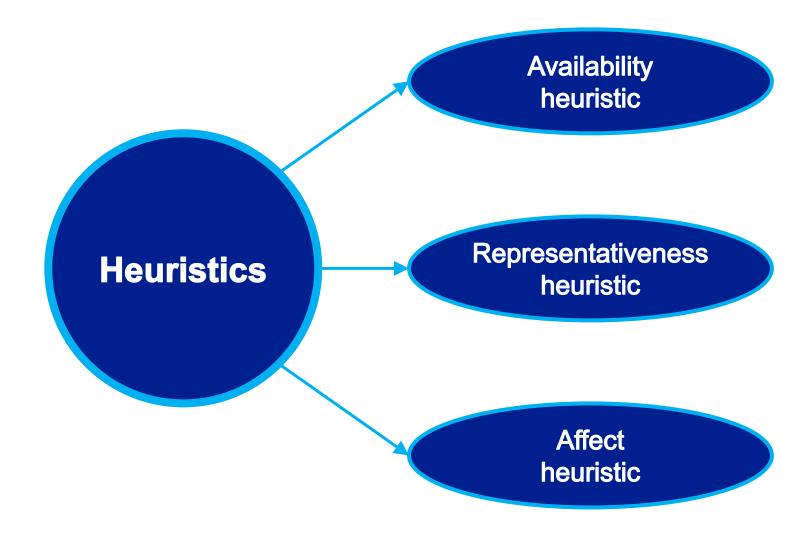


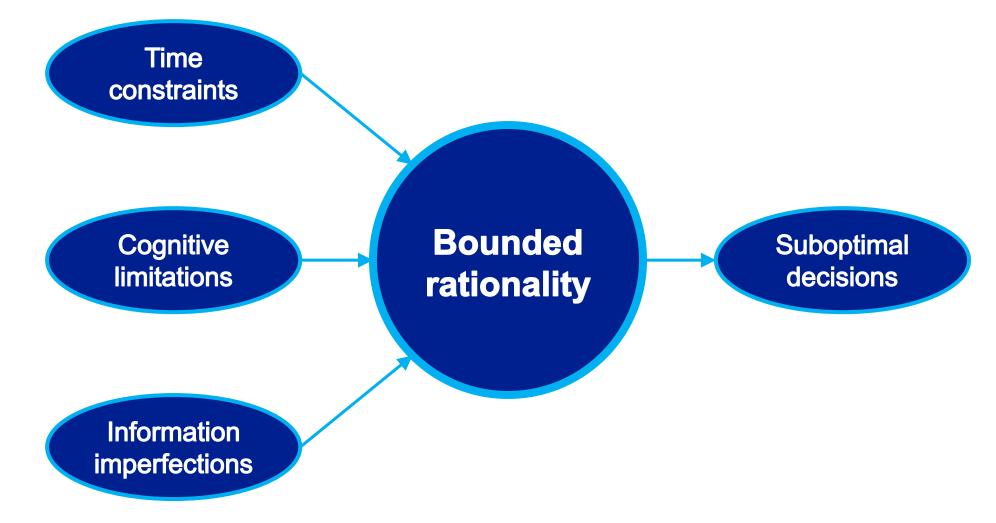
WIPO FOR OFFICIAL USE ONLY

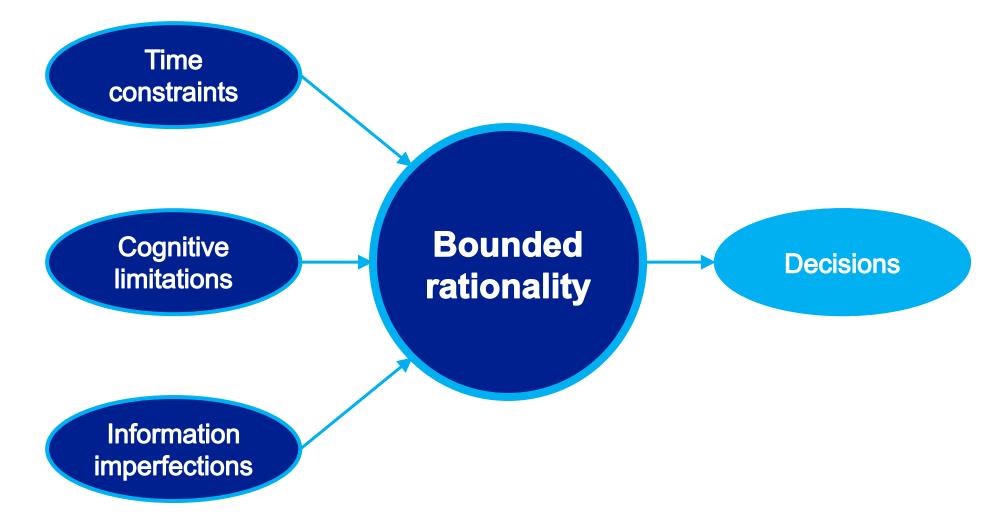
## Why the problem exists?











# How do we try to communicate effectively?

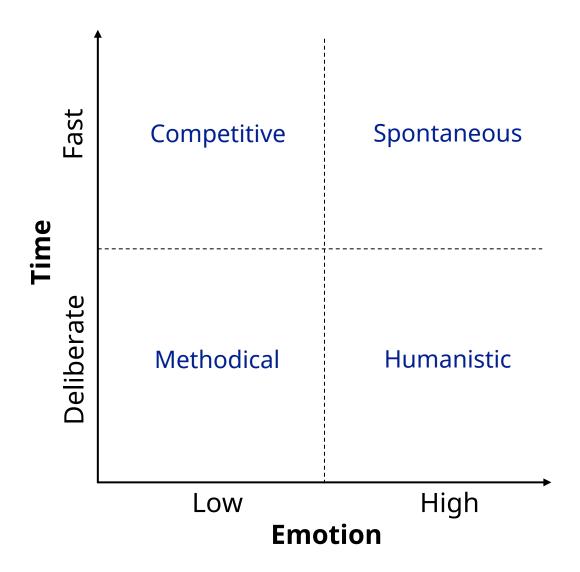
- The importance of IP education
- The question behind the question
- The power of advanced analytics software
- The role that visualizations play
- Data storytelling
- The feedback loop

Policymakers want to engage with the decision...

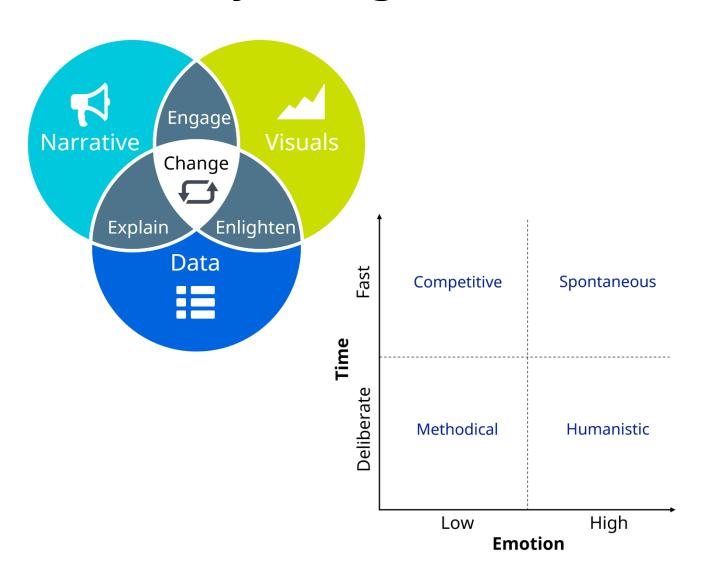
not necessarily the data



## Effective decision making



## Data storytelling



People hear statistics, but feel stories



AI is changing the role of the patent analyst, so embrace it

How you communicate your analysis is key; people hear statistics, but feel stories

## Patent analytics at WIPO





#### **WIPO patent landscape reports**

- Over 15 PLR reports on various topics prepared since 2010
- Compilation of publicly available PLRs



45

## Patent analytics methodological resources

- WIPO guidelines for preparing patent landscape reports
- Manual on free and opensource tools for patent analytics
- Handbook on patent analytics



#### **WIPO technology trends**

 Big-scale patent landscaping exercise with additional data analysis, contextualization and expert inputs



WIPO FOR OFFICIAL USE ONLY

#### **Patent analytics training**

 Introductory and advanced training webinars/seminars and workshops







© WIPO, 2023



Attribution 4.0 International (CC BY 4.0)

The CC license does not apply to non-WIPO content in this presentation.

Photo credits: WIPO, B.Rossen, K. Lurie, L. Wretblad, Alpha Stock Images

# Thank you for listening!

christopher.harrison@wipo.int

IP.analytics@wipo.int