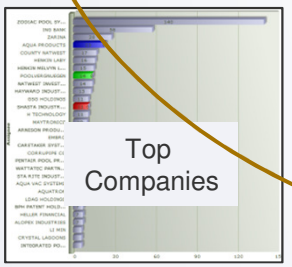
Customizing Statistics for Sharper Analysis

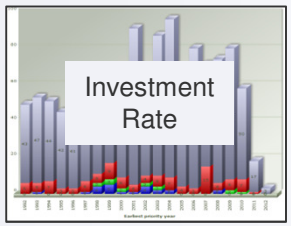
Laurent Hill
Renaud Garat

April 15, 2013

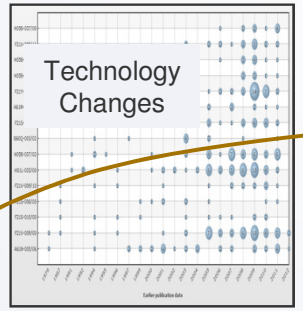
Conceptual Model Of A Patent Based Analysis

1. Understand the Overall Landscape

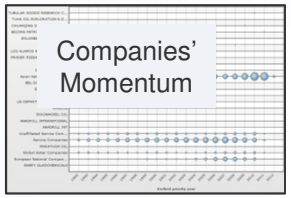


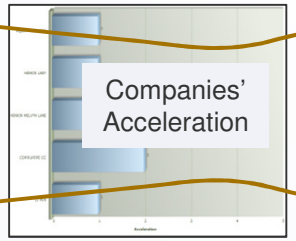


2. Understand Recent Trends

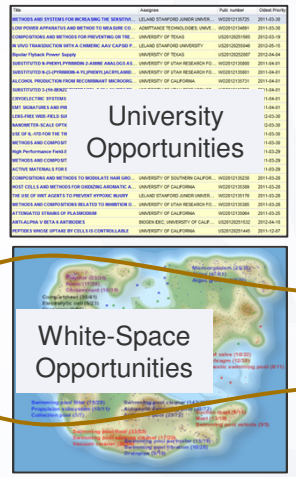


3. Understand Close Art

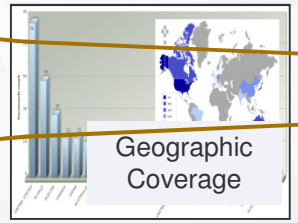


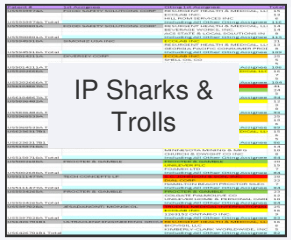


4. Understand Available Art

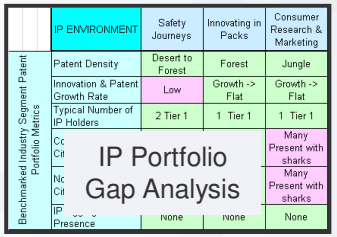


5. Estimate Freedom To Practice

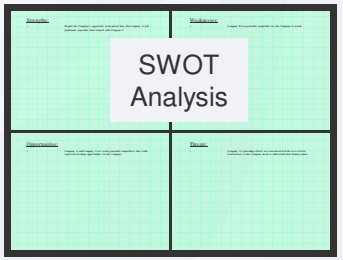




6. Determine Who To Watch And Leverage



Strategic Direction



Data → Information → Knowledge + Experience → Insight

Why Customize Analysis?

- Data quality
- Productivity
- Technical Specific Needs

Why Customize Analysis?

- Data quality
  - Names: normalisation – M&A

    bank collaterals
  - Concepts: normalisation

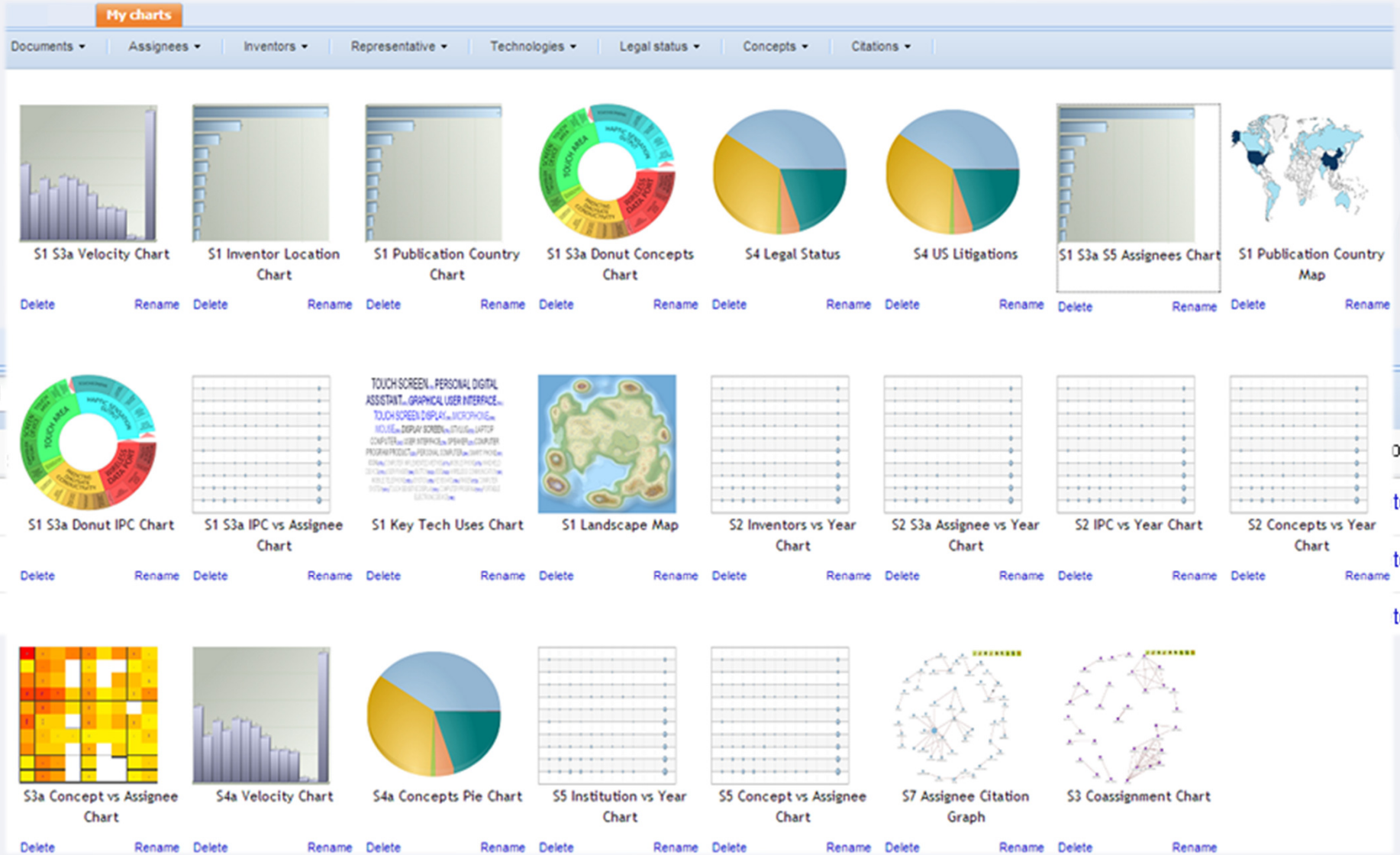
    languages

Assignee
EXXONMOBIL RESEARCH ENGINEERING
EXXONMOBIL CHEMICAL PATENTS
EXXONMOBIL PRODUCTION RESEARCH
EXXONMOBIL
EXXONMOBIL OIL
EXXONMOBIL UPSTREAM RESEARCH
INFINEUM
BANK OF AMERICA
CALGON
CITICORP
DATAPRODUCTS
IMAGING SOLUTIONS
NALCO
NALCO CROSSBOW WATER
NALCO ONE SOURCE
ADVANCED ELASTOMER SYSTEMS
EXXON NUCLEAR

Why Customize Analysis?

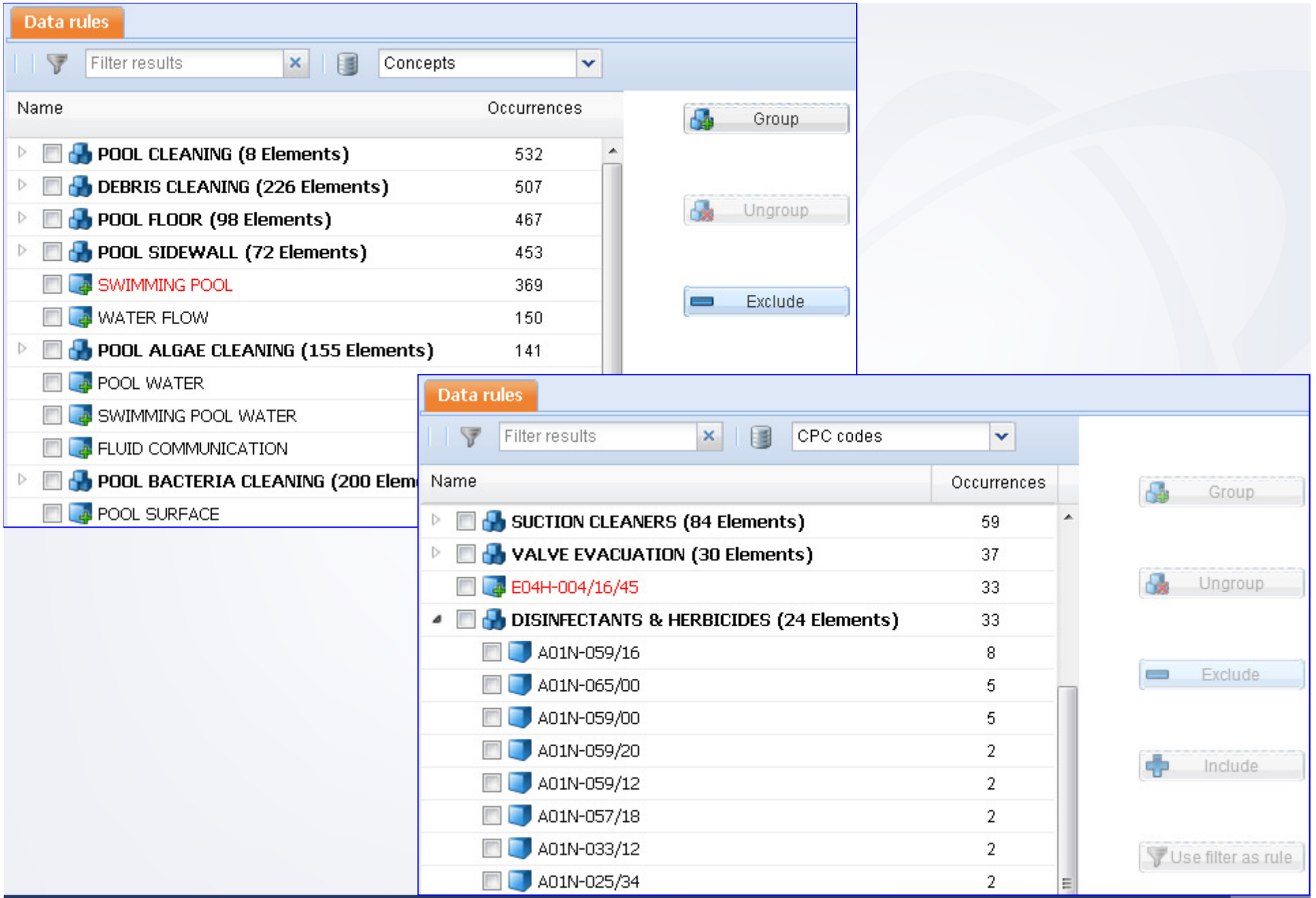
- Productivity
  - Guideline with predefined set of charts
  - Re-usable templates:

    charts, axis, colors, grouping rules (assignees, concepts, classification codes, inventors…), real-time interaction (refining/expanding)


My charts
Documents
Assignees
Inventors
Representative
Technologies
Legal status
Concepts
Citations
S1 S3a Velocity Chart
S1 Inventor Location Chart
S1 Publication Country Chart
S1 S3a Donut Concepts Chart
S4 Legal Status
S4 US Litigations
S1 S3a S5 Assignees Chart
S1 Publication Country Map
S1 S3a Donut IPC Chart
S1 S3a IPC vs Assignee Chart
S1 Key Tech Uses Chart
S1 Landscape Map
S2 Inventors vs Year Chart
S2 S3a Assignee vs Year Chart
S2 IPC vs Year Chart
S2 Concepts vs Year Chart
S3a Concept vs Assignee Chart
S4a Velocity Chart
S4a Concepts Pie Chart
S5 Institution vs Year Chart
S5 Concept vs Assignee Chart
S7 Assignee Citation Graph
S3 Coassignment Chart
Delete
Rename

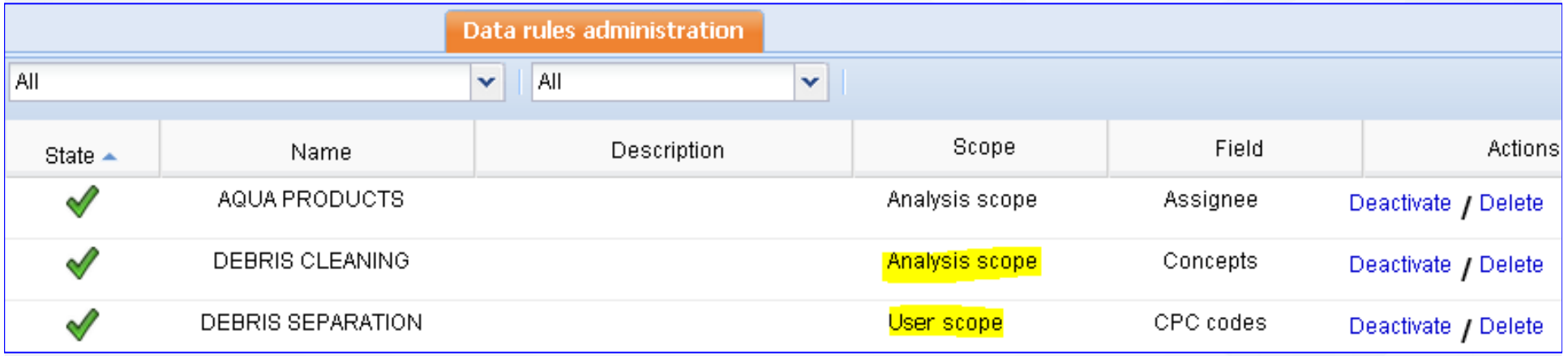
Why Customize Analysis?

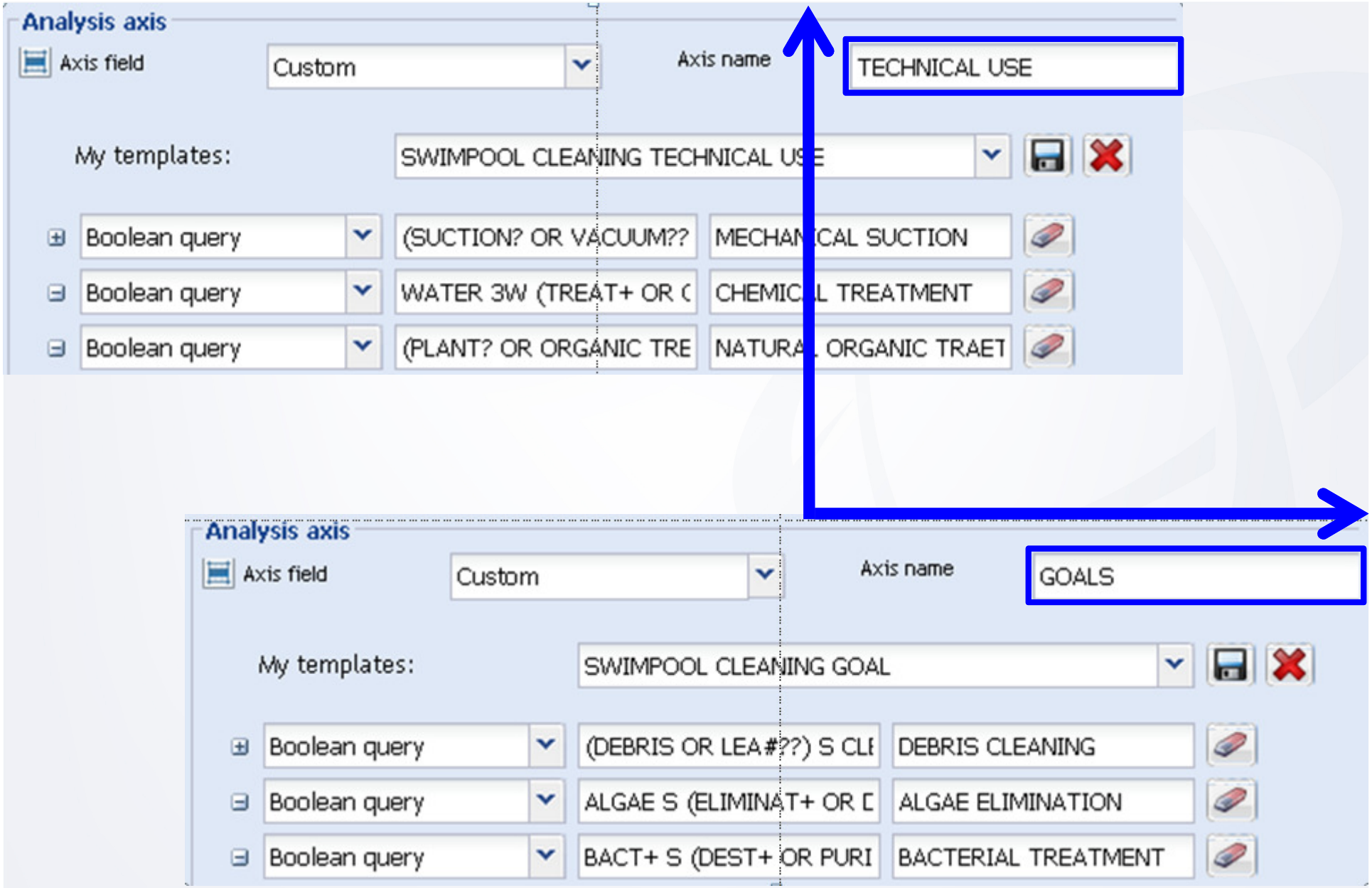
- Technical Specific Needs
  - Concepts grouping / hiding
  - Classification codes grouping / hiding


Data rules
Filter results
Concepts
Name
Occurrences
POOL CLEANING (8 Elements) 532
DEBRIS CLEANING (226 Elements) 507
POOL FLOOR (98 Elements) 467
POOL SIDEWALL (72 Elements) 453
SWIMMING POOL 369
WATER FLOW 150
POOL ALGAE CLEANING (155 Elements) 141
POOL WATER
SWIMMING POOL WATER
FLUID COMMUNICATION
POOL BACTERIA CLEANING (200 Elem
POOL SURFACE
Group
Ungroup
Exclude
Data rules
Filter results
CPC codes
Name
Occurrences
SUCTION CLEANERS (84 Elements) 59
VALVE EVACUATION (30 Elements) 37
E04H-004/16/45 33
DISINFECTANTS & HERBICIDES (24 Elements) 33
A01N-059/16 8
A01N-065/00 5
A01N-059/00 5
A01N-059/20 2
A01N-059/12 2
A01N-057/18 2
A01N-033/12 2
A01N-025/34 2
Group
Ungroup
Exclude
Include
Use filter as rule

Why Customize Analysis?

- Problems to solve
  - Grouping Rules: Permanent vs Contextual




Analysis axis
Axis field
Custom
Axis name
TECHNICAL USE
My templates:
SWIMPOOL CLEANING TECHNICAL US
Boolean query
(SUCTION? OR VACUUM??
MECHA CAL SUCTION
Boolean query
WATER 3W (TREAT+ OR (
CHEMIC L TREATMENT
Boolean query
(PLANT? OR ORGANIC TRE
NATURA ORGANIC TRAET
Analysis axis
Axis field
Custom
Axis name
GOALS
My templates:
SWIMPOOL CLEANING GOAL
Boolean query
(DEBRIS OR LEA#??) S CL
DEBRIS CLEANING
Boolean query
ALGAE S (ELIMINAT+ OR D
ALGAE ELIMINATION
Boolean query
BACT+ S (DEST+ OR PURI
BACTERIAL TREATMENT

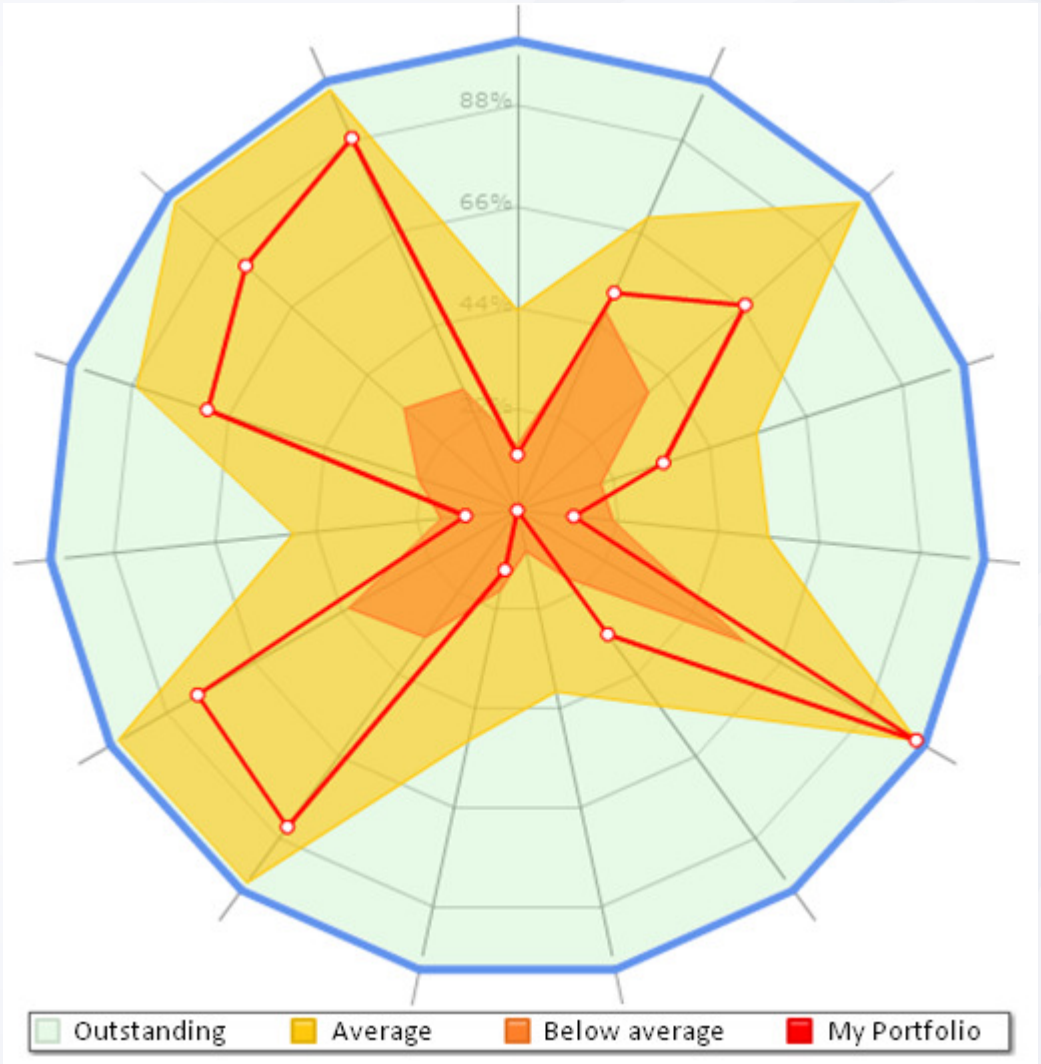
Why Customize Analysis?

- Conclusion
  - ☺ Powerful and precise
  - ☹ Time consuming (system programing?)
  -  Future…

Future…

Guided activities

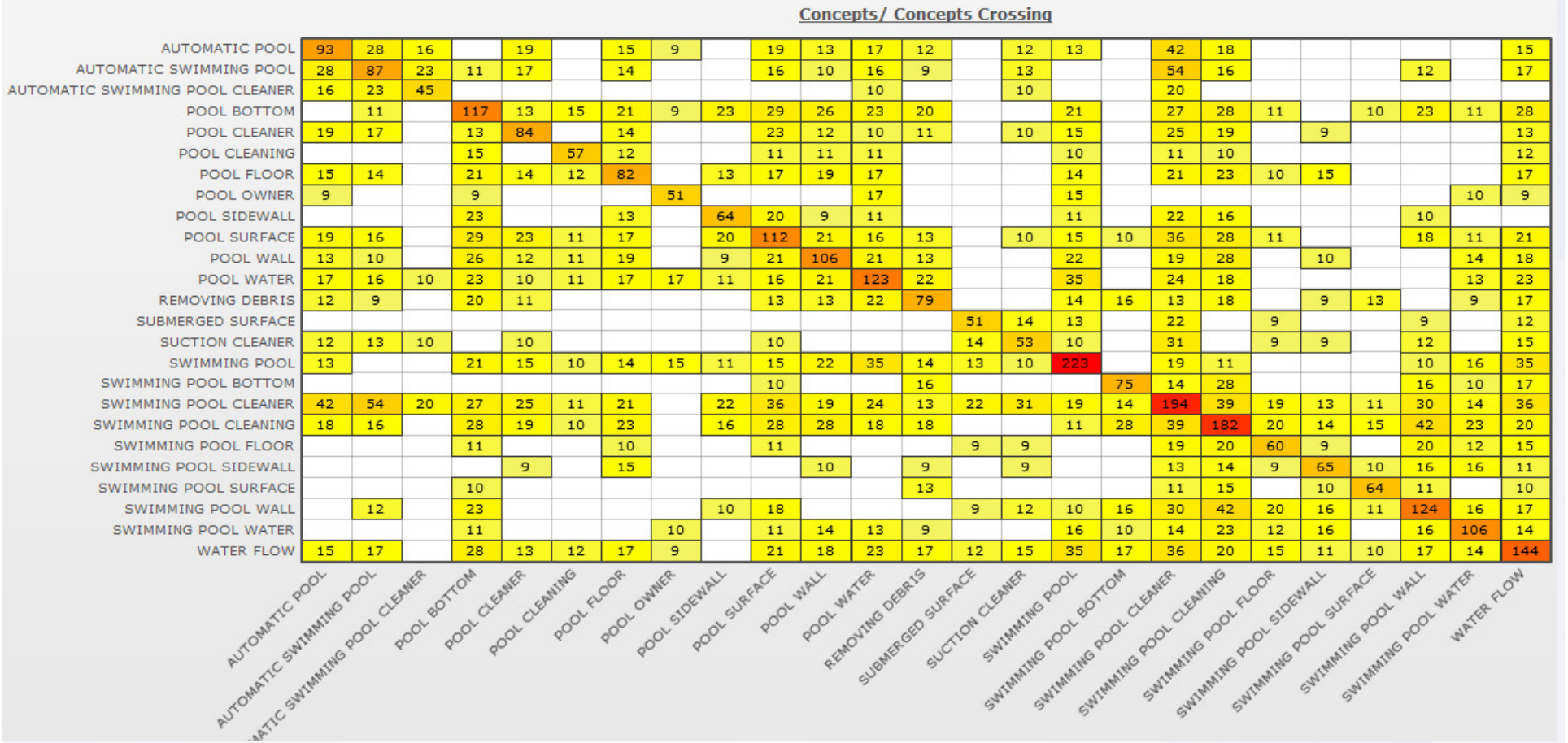
Preconfigured Analysis Metrics



Case study: building a technology/use matrix

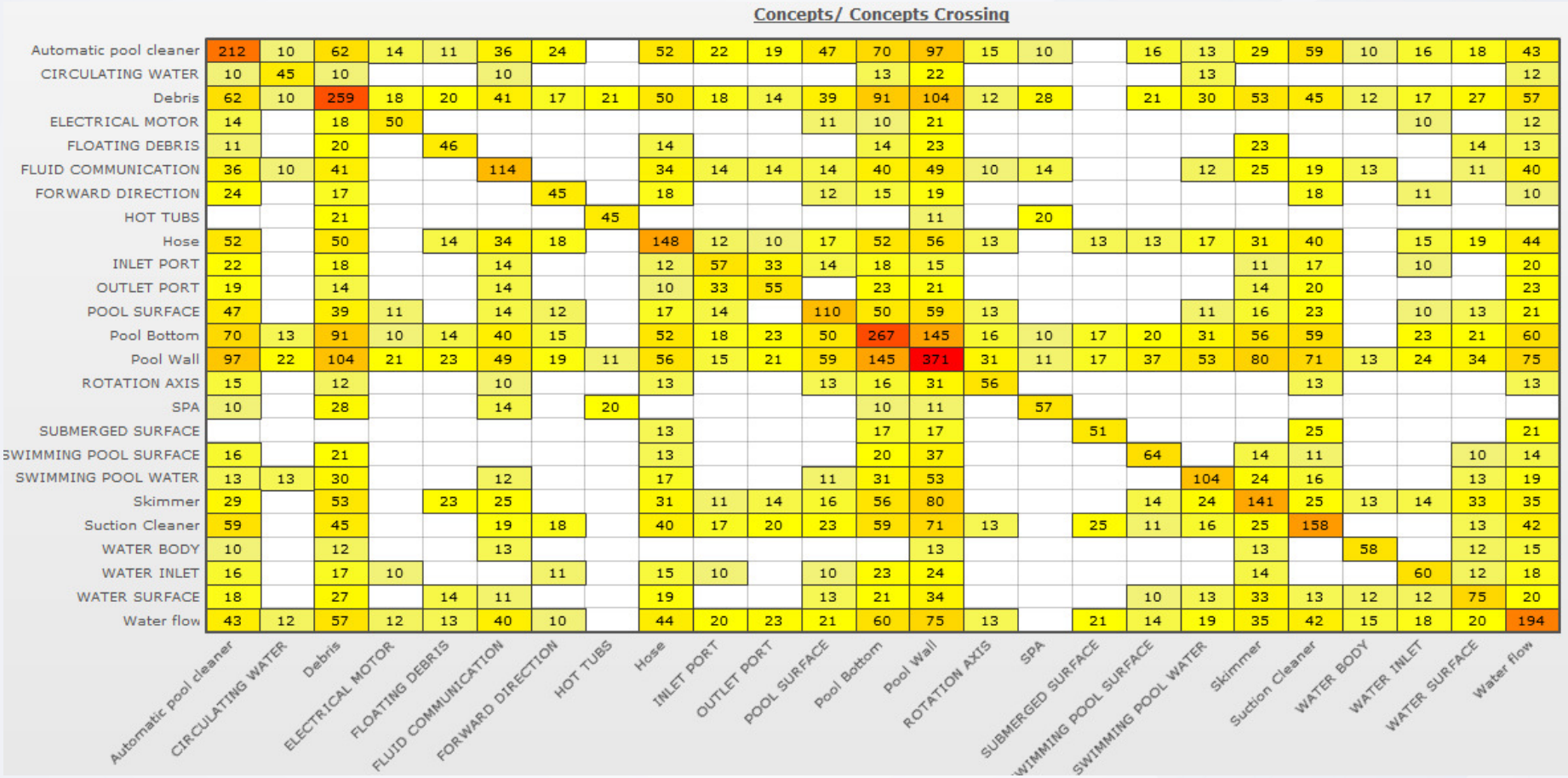
- The goal is to understand which technologies are used to achieve a given goal
- Corpus : Swimming pool cleaners
  - ~ 1100 patents
- Let's explore a few alternatives on how to do it

First try: concept/concept matrix



- No clear picture emerges, the concepts just reflect important notions in the corpus

Concepts after clean up / Data rules



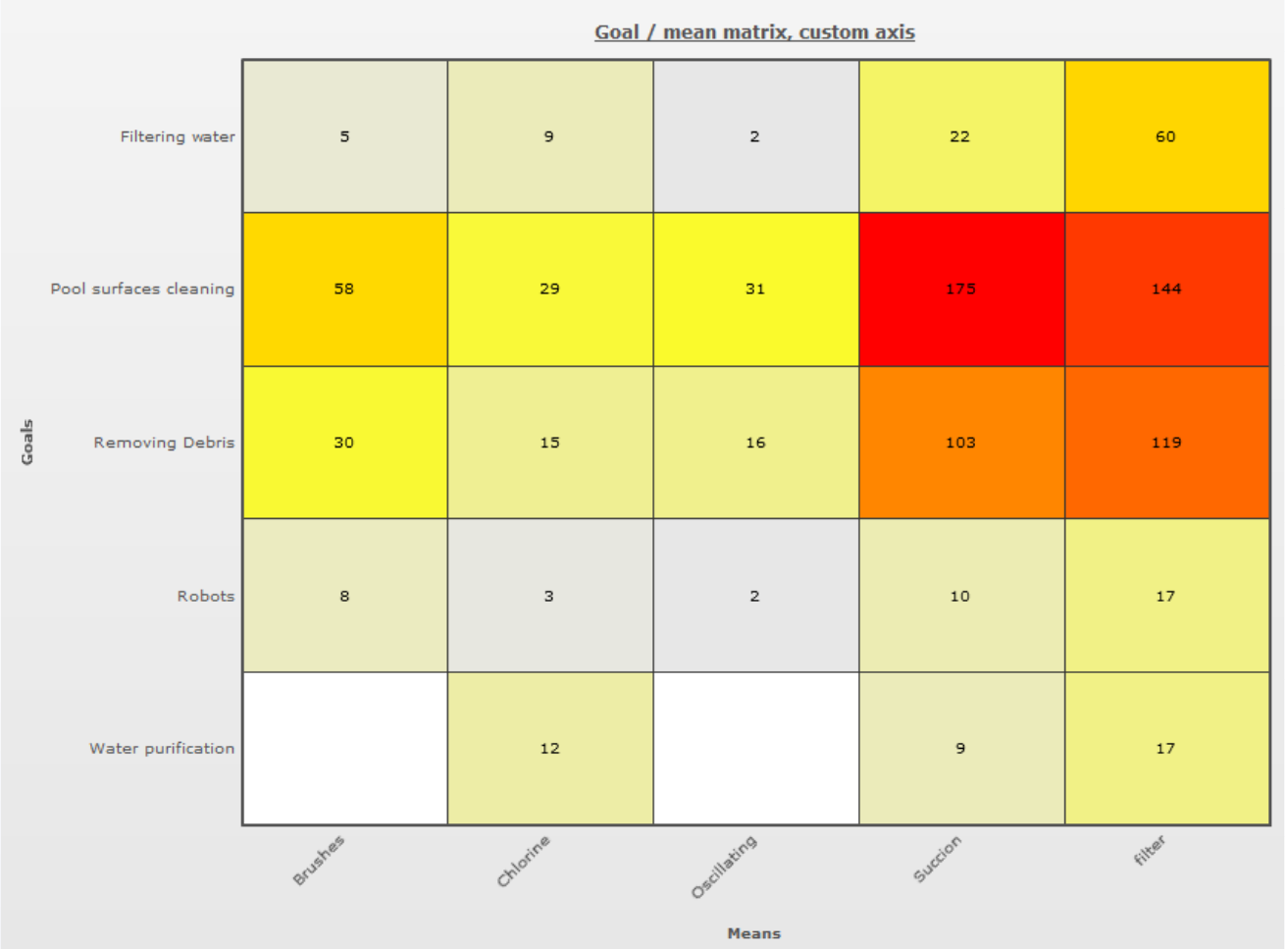
- Getting better, drilling down on each cell very powerful for understanding

We need a better solution

- Clearly this is not enough, we need to separate problem domain concepts from solution concepts
- Solution: custom axis
  - Let's start by manually sorting out concepts
  - One axis for Goals , one axis for Means

Goal / mean matrix, concept version

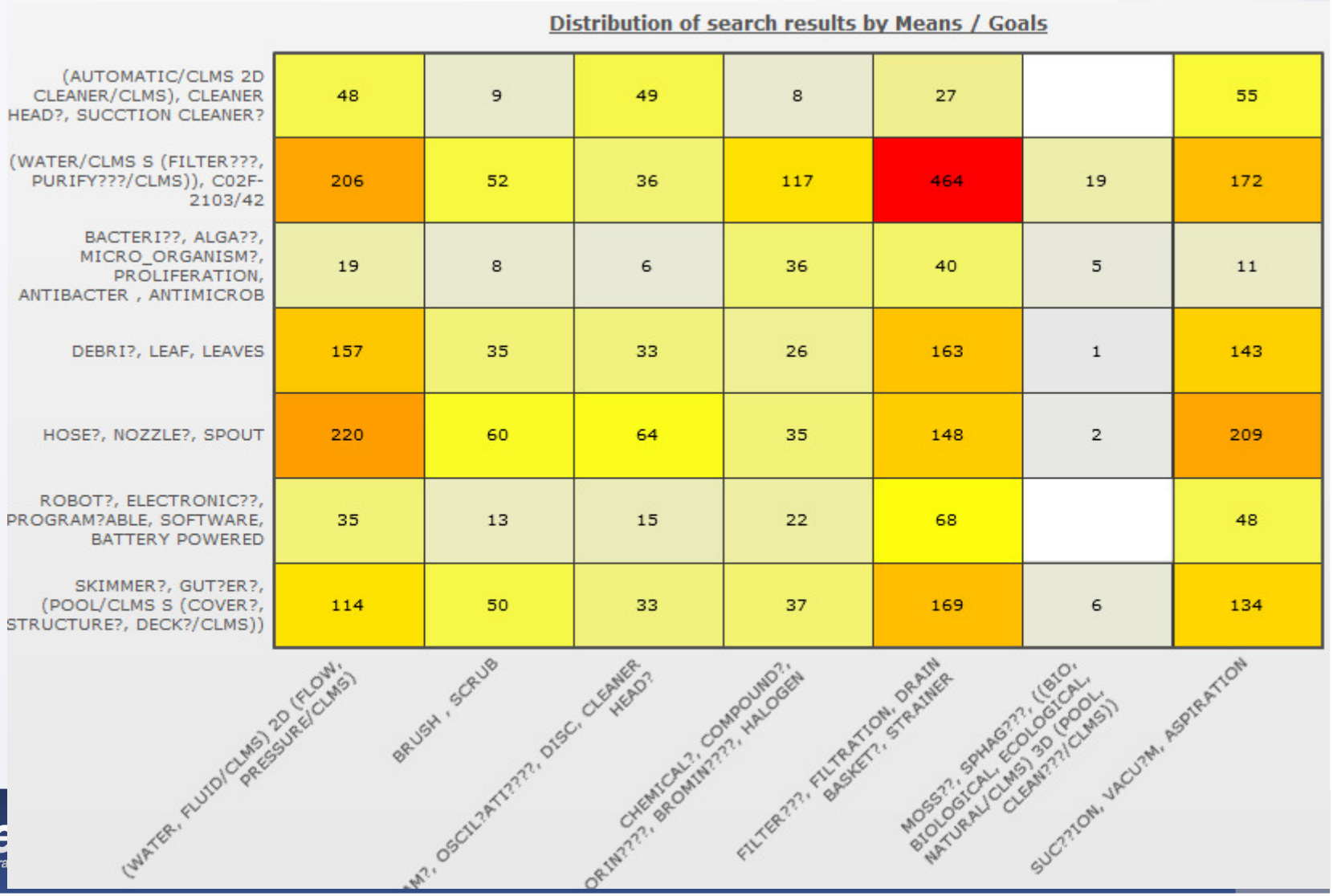
- Simple and clear:
  - Good for communication
- Precision and completeness could be improved



Slice & Dice through boolean queries

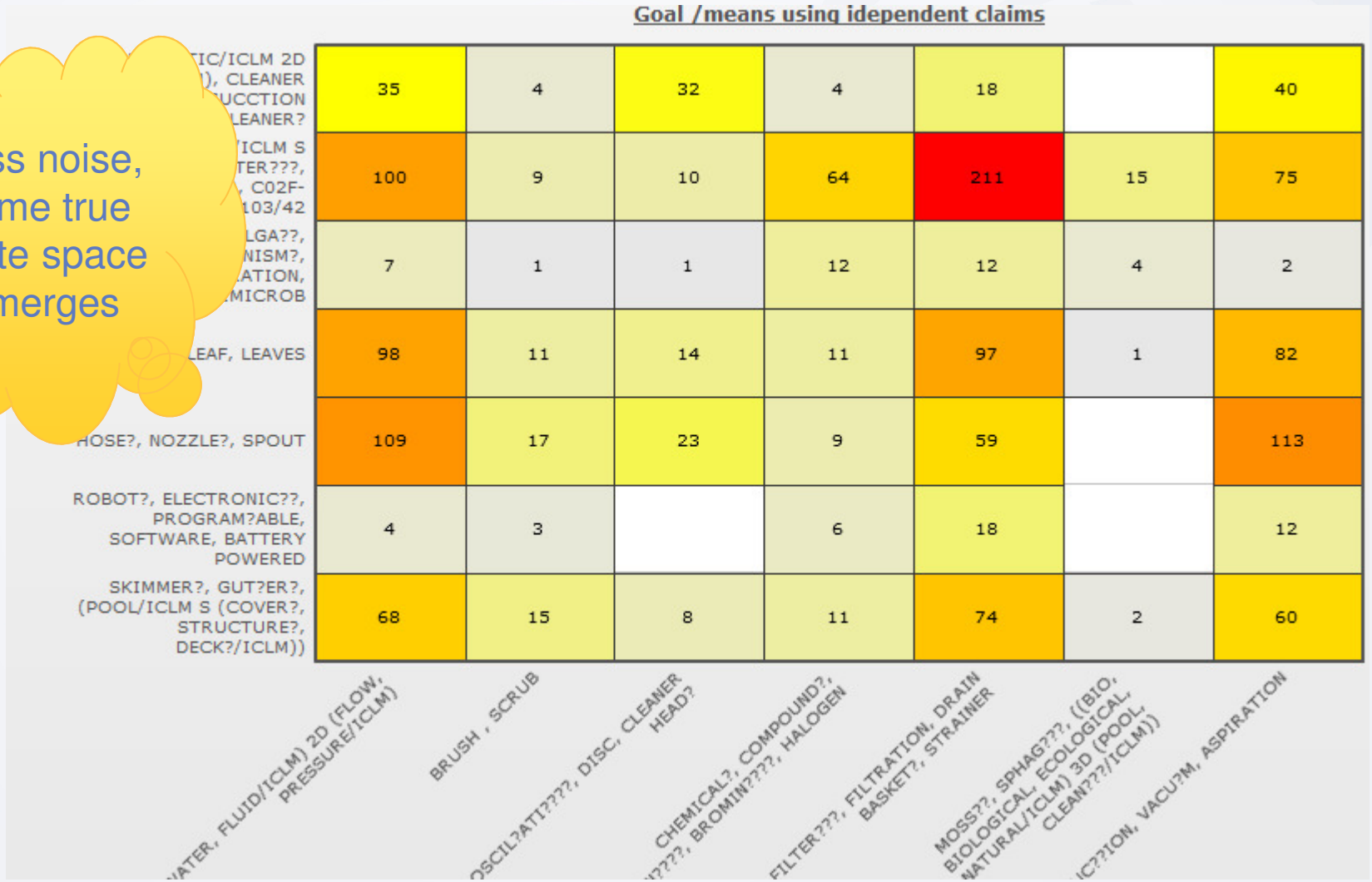
- We need a sharper tool to dissect the corpus
- We already have a sharp tool with the full text search engine: fields, proximity…
- Each value in a custom axis can be a search engine query.
- Example: combining independent claims and CPC
  - ((water 3D (filter??? or purify???))/ICLM or C02F-2103/42/CPC)

Technology crossing from Custom analysis axis



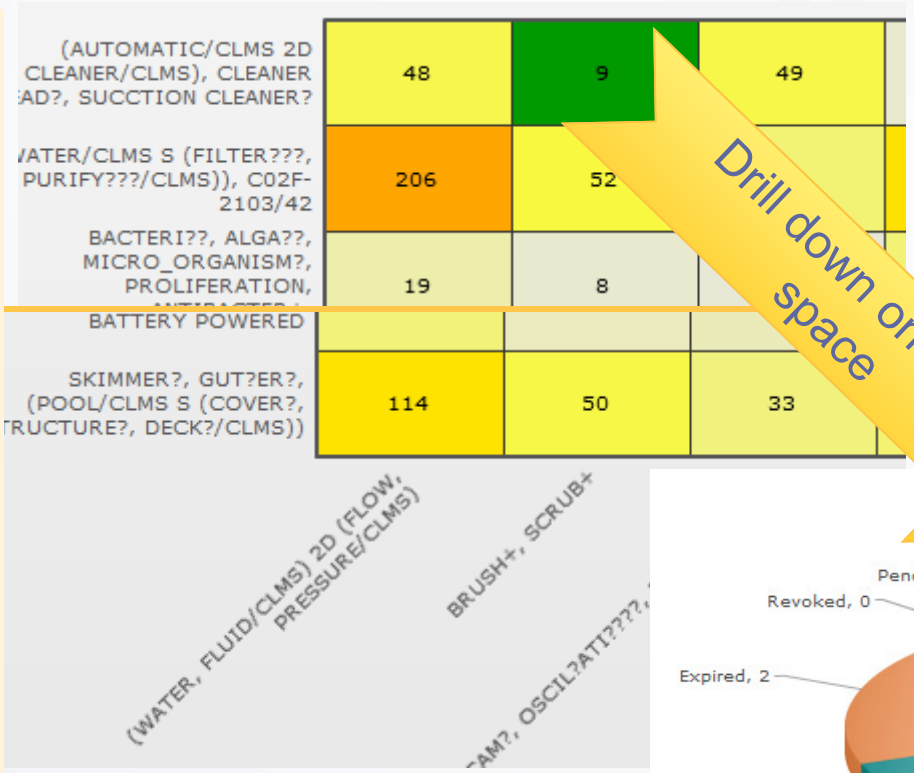
Another view, using independent claims

Less noise, some true white space emerges

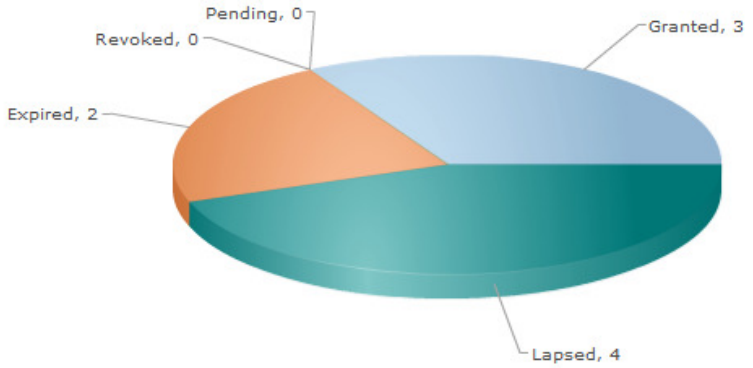


Analyzing white space on technology crossings

- No pending patents
- Most of the art is lapsed or expired
- Dead end?



Drill down on white space



Case study wrap up

- Editing data, customizing analysis views and axis let us go beyond simple analysis
- By leveraging the precision of a full featured search engine on all patent text, we were able to build a goals/Mean technology matrix
- Domain expertise still needed, no magic bullet

Thank you

Laurent Hill
Renaud Garat
April 15, 2013