

From Patstat to TIP

The logical path of a patent landscape analysis: from Espacenet to Patstat, from Patstat to TIP

Riccardo Priore, EPIP Conference 2026

What Is a Patent Landscape Analysis?

«A patent landscape report is an analytical tool that provides an overview of the patent situation relating to a particular technology, in a particular country or region, or for a particular assignee.»

WIPO, Patent Landscape Reports (publication series page)

The Journey in Three Steps

- ▶ Foreword — Espacenet: the starting point, public and accessible to anyone, but with practical limits of scale
- ▶ Part I — Patstat queried in SQL: overcomes Espacenet's limits and, on its own, already covers most of what a credible PLR needs
- ▶ Part II — TIP: the same underlying data, extended with Python and web tools within the same working session
- ▶ The common thread: from raw data to a narrative a decision-maker can act on



FOREWORD

Espacenet: The Starting Point

What Espacenet Is

- ▶ The EPO's free, publicly available patent search engine, accessible via browser with no licence required
- ▶ No SQL or programming skill required: designed to be usable from day one, even by a complete beginner
- ▶ Near-worldwide coverage of patent documents, with simple, intuitive search forms
- ▶ For the large majority of analysts, it is the first tool through which contact is made with patent data

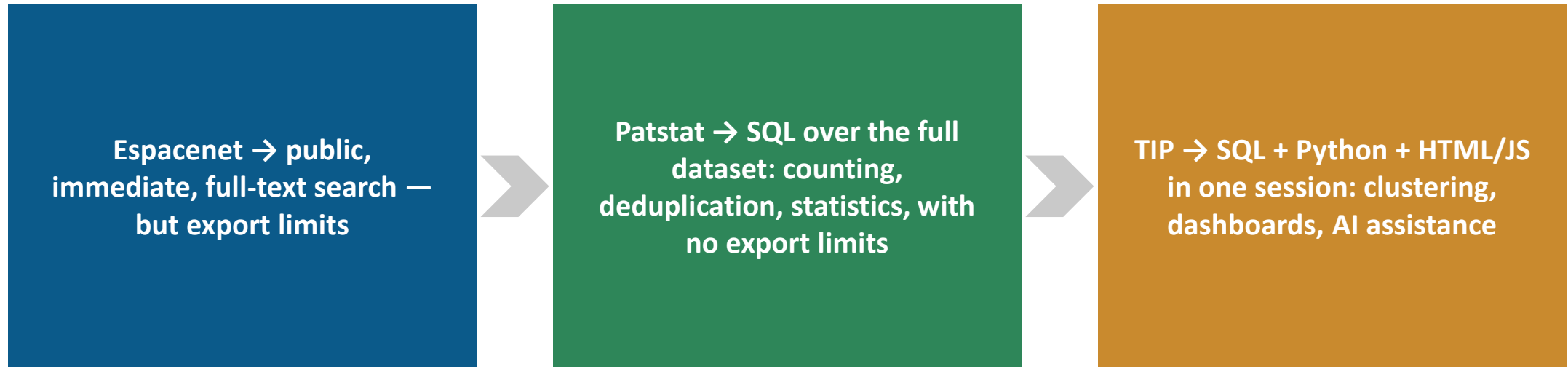
Espacenet's Strength: Full-Text Search

- ▶ Unlike Patstat and TIP, Espacenet allows direct full-text and claims searching
- ▶ Particularly useful in the early stages, when documents must be identified from specific textual wording, not only from classification codes
- ▶ A capability neither Patstat nor TIP natively replicates in the same way, and a reason to return to Espacenet even after moving on to more powerful tools

Espacenet's Limitation: Scale

- ▶ Results can only be exported up to a maximum number of records per session
- ▶ For a broad landscape, this forces multiple, repeated download sessions
- ▶ With very large result sets, the operation sometimes becomes practically unfeasible, or at best discouraging
- ▶ No family deduplication, no server-side aggregate statistics, no large-scale relational querying: systematic analysis requires a different tool

The Complete Path, in Brief



Each step exists to overcome the practical limit of the one before it

PART I

From Espacenet to Patstat: What SQL Can Already Do

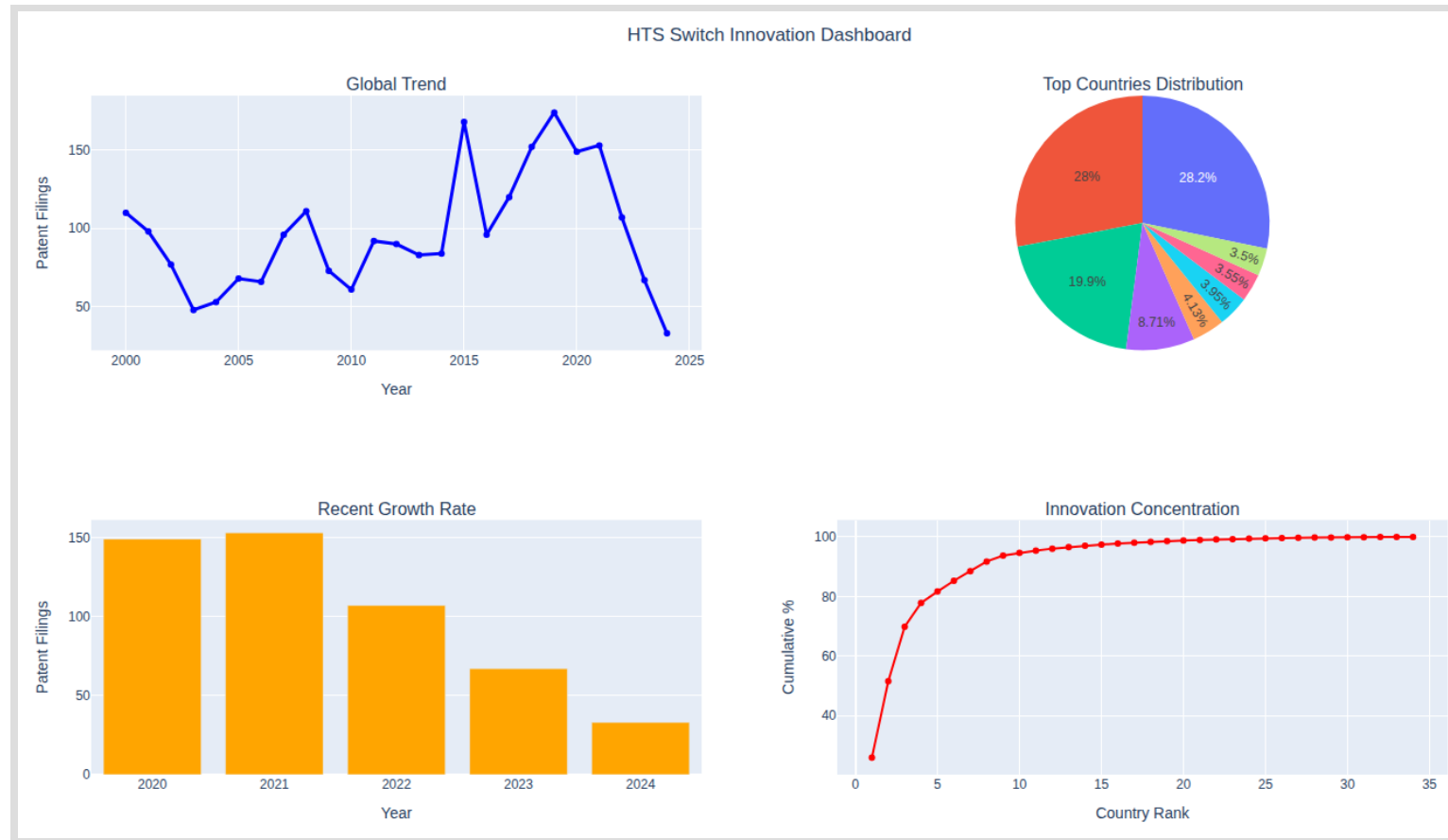
The Basic Building Blocks

- ▶ Counting without double-counting: patent-family deduplication
- ▶ Technology classification via IPC/CPC codes
- ▶ Co-occurrence matrices between codes, mapping the relationships between technologies
- ▶ Equivalent families (docdb_family_id): finding what keywords alone cannot

Who, Where, and Who Cites Whom

- ▶ Ranking applicants and inventors; geographic distribution of families
- ▶ Backward and forward citations: where an invention comes from, how much it influenced what follows
- ▶ Time trends, averages, contingency tables (CASE WHEN)

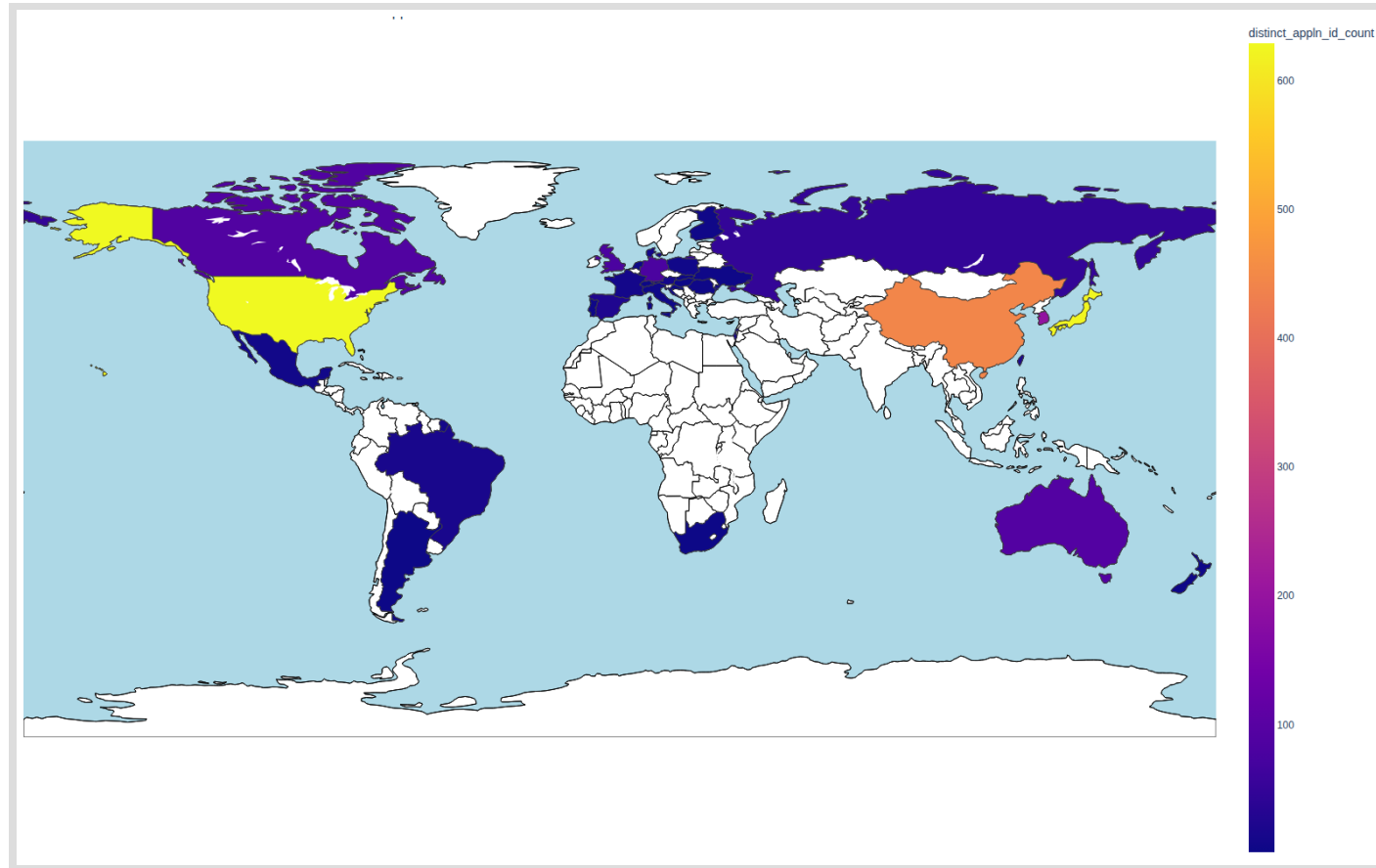
From Tabular Data to a First Chart



Summary dashboard built directly from SQL counts: trend, distribution, geographic concentration

Source: DEEPTeCH-Russo, HTS Switches Innovation Dashboard

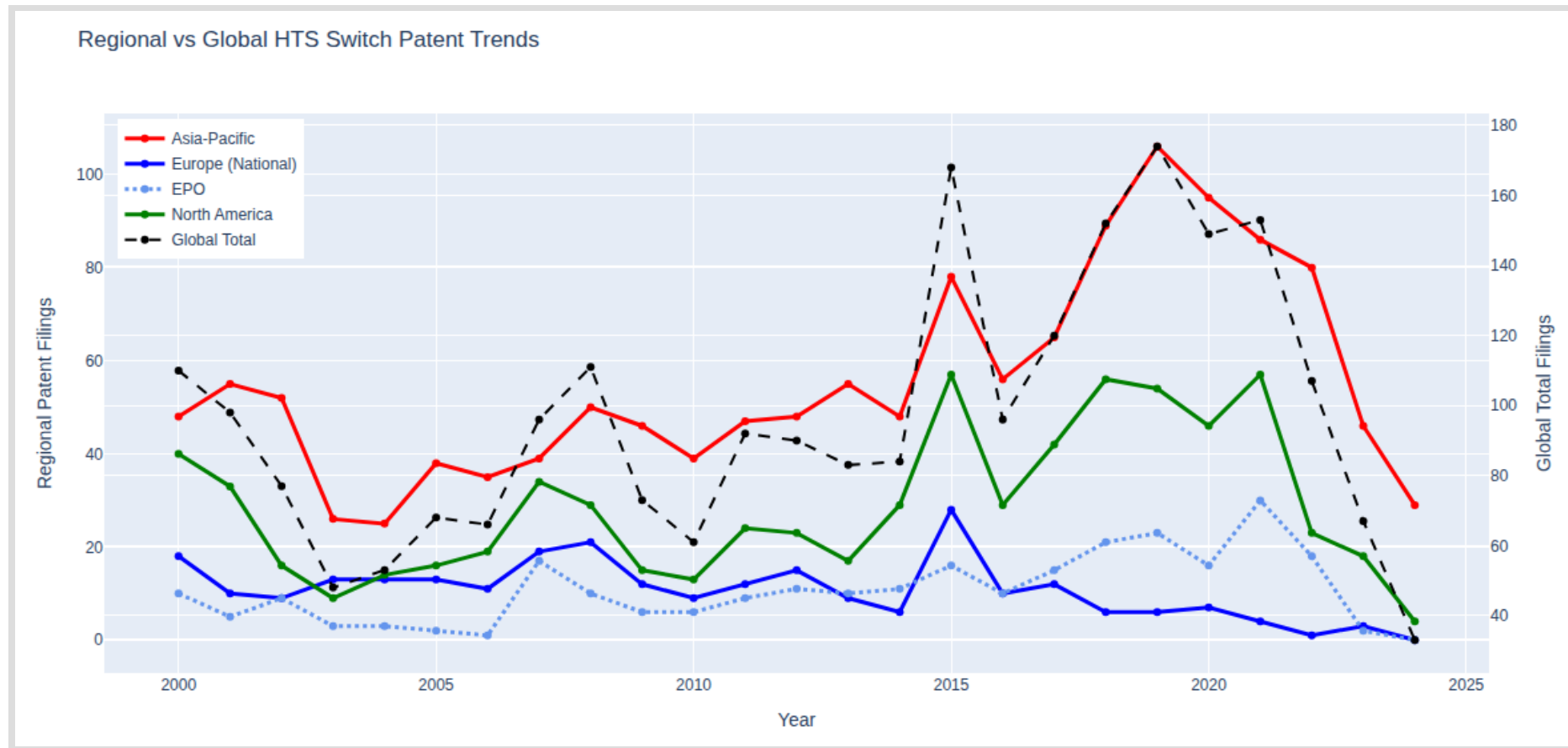
The Geography of a Technology



Territorial coverage of patent families, by distinct application count

Source: DEEPTeCH-Russo, HTS Switches Filings Map

Telling a Story Over Time



Regional vs. global trend: the same SQL query, read over time, becomes a narrative

Source: DEEPTech-Russo, Regional vs Global HTS Switch Trends

Quality and Legal Status

- ▶ Triadic families and renewal-fee payments as proxies for economic value
- ▶ EP Register: examination citations, licences, ownership changes, oppositions and revocations
- ▶ Text mining (CONTAINS/FREETEXT) and the first built-in graphic tools

Where SQL Alone Reaches Its Limit

- ▶ No native clustering, dimensionality reduction, or machine learning (e.g. t-SNE)
- ▶ Inherently tabular output: no interactive dashboards
- ▶ Any further processing requires exporting to another tool, breaking the continuity of the analysis

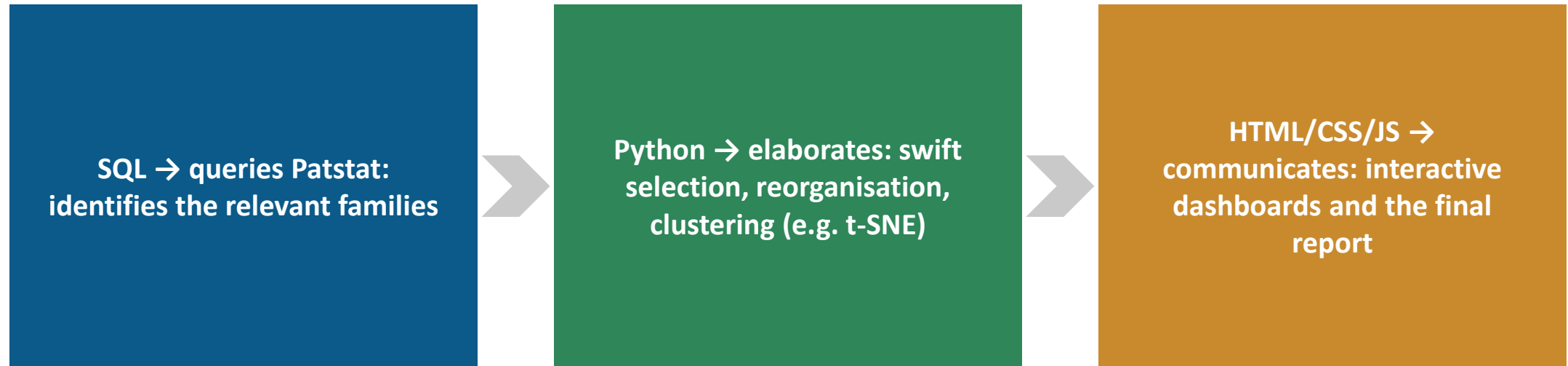
PART II

Beyond Patstat: Extending the Analysis with TIP

What TIP Is

- ▶ The EPO's Technology Intelligence Platform: JupyterLab in the cloud, tip.epo.org
- ▶ Direct access to Patstat via epo.tipdata.patstat (PatstatClient, SQLAlchemy ORM)
- ▶ The same tables as the HANDBOOK, now queryable in SQL or Python within the same session
- ▶ Beyond widening the analytical perspective, TIP's operational speed matters: less time spent on export/import cycles frees the analyst to focus on refining the search strategy itself

One Session, Many Languages



All within the same notebook: no intermediate exports

Same Data, Richer Interaction

FlowNano -- Triadic Patent Families: Country of Residence x Applicants
For each country of residence detectable among triadic-family applicants, the corresponding applicants and their triadic-family counts.

611
Triadic Families

33
Countries Detected

920
Unique Applicants

All countries
Search applicant name...
937 row(s) shown

Country	Country Families	Applicant	N Triadic Families	Family IDs (hover for full list)
AT	2	HARRER, MARTIN	1	47469991
AT	2	WHITEHEAD, ADAM	1	47469991
AT	2	GILMEISTER ENERGY STORAGE	1	47469991
AT	2	CELLSTROM	1	47469991
AT	2	DICHAND, MICHAEL	1	48790499
AU	13	REDFLOW R&D	7	39787954; 41706774; 43899739; 52279222; 52827444; 53799452; 54008060
AU	13	REDFLOW	3	39787954; 41706774; 43899739
AU	13	WINTER, ALEXANDER RUDOLF	3	39787954; 41706774; 43899739
AU	13	NEWSOUTH INNOVATIONS	2	37570018; 70461808
AU	13	GELION TECHNOLOGIES	2	53777059; 69776458
AU	13	KAZACOS, NICHOLAS, MICHAEL	1	37570018
AU	13	KAZACOS, GEORGE, CHRISTOPHER	1	37570018
AU	13	V-FUEL	1	37570018
AU	13	COAD, NATHAN	1	47746829
AU	13	GELION TECHNOLOGIES PTY LTD.	1	53777059
AU	13	UNIVERSITY OF SYDNEY	1	53777059
AU	13	MONASH UNIVERSITY	1	56012948
AU	13	GELION TECHNOLOGIES PTY LTD	1	69776458
AU	13	NEWSOUTH INNOVATIONS, PTY LIMITED	1	70461808
BE	4	SOLVAY	2	43481627; 48652120
BE	4	VANDENBORRE, HUGO	1	43481627

Source: PATSTAT (EPO TIP Production) | Period: 2000-2024 | Four-branch IPC+keyword search (cross-pollination rule)

Explorable, filterable table built on Patstat data extracted in TIP: 611 triadic families, 920 applicants

Source: DEEPTTECH-FlowNano, Country x Applicants Explorer

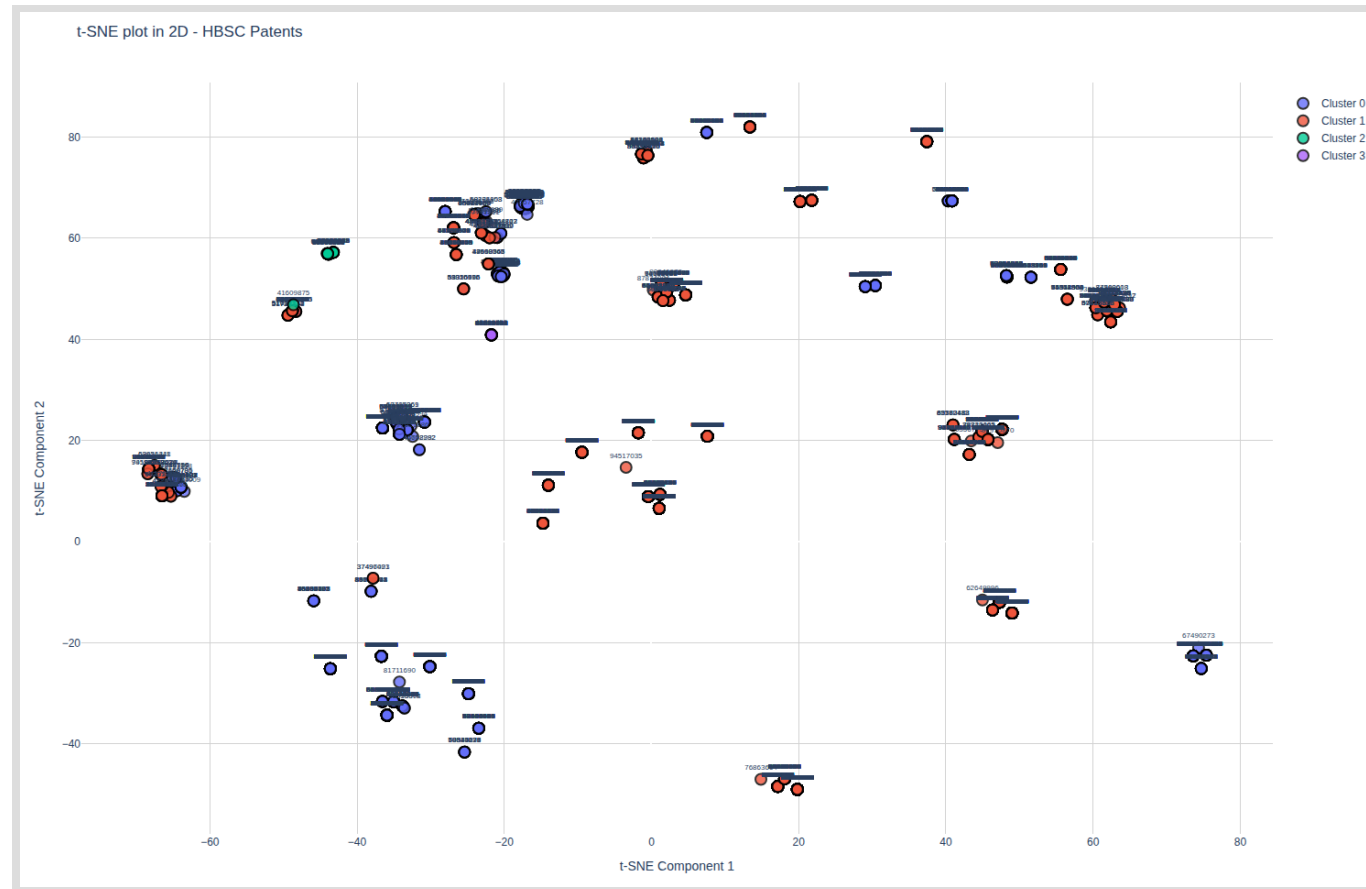
Clustering Becomes Explorable



Interactive dashboard based on t-SNE + K-Means: 7,795 families grouped into 4 technology clusters

Source: DEEPTech-Novac, HBSC Patent Landscape Cluster Explorer

A Concrete Example: The t-SNE Cited by WIPO

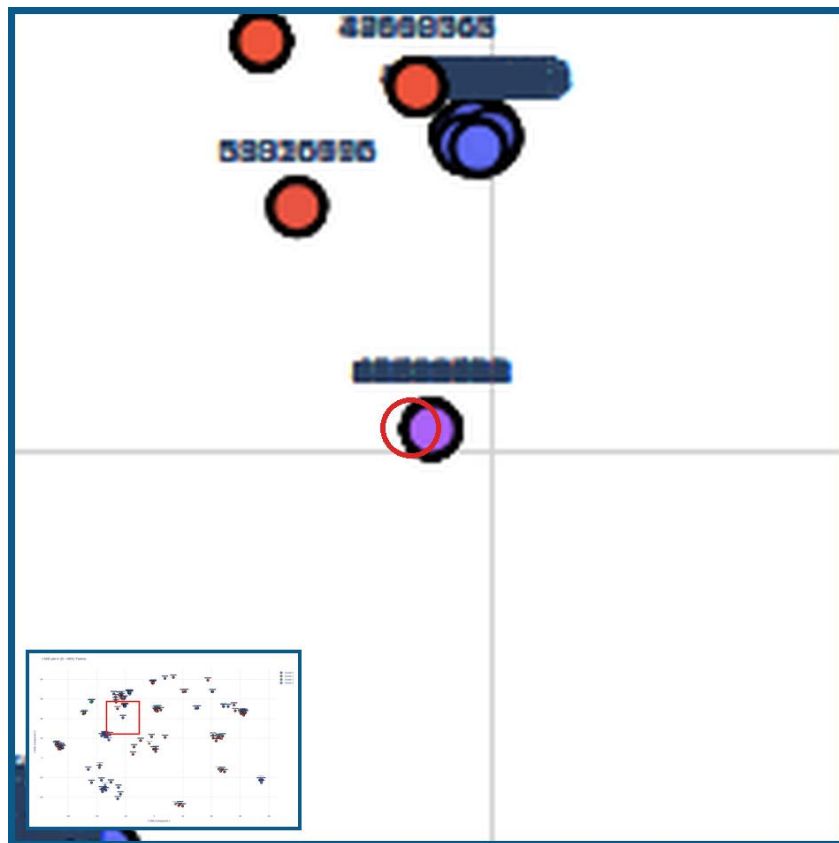


Dimensionality reduction of IPC codes: technological groupings invisible to a single SQL query

Source: DEEPTeCH-Novac, HBSC t-SNE Analysis (2D)

From Cluster to Document: Full Granularity

ON THE T-SNE PLOT



Cluster 3 — Battery/Supercapacitor Charging Systems: 16 patent families collapse into a single point. The overlapping label is unreadable in a static image.

EXPLODED INSIDE THE SAME TIP SESSION

Publication No.	Title	Applicant
WO2022091673A1 (2020)	Abnormality Detection Method, Abnormality Detection Device, Power Storage Device, and Computer Program	GS Yuasa International Ltd.
EP3268974A1 (2015)	Nickel Supercapacitor Engine Starting Module	Saft America
US2013162216A1 (2011)	Stacks of Internally Connected Surface-Mediated Cells and Methods of Operating Same	Zhamu, Aruna
EP3510687A1 (2016)	Supercapacitor Charge System and Method	E-Synergy Graphene Research Pte. Ltd.
US11340299B2 (2017)	Systems and Methods for Monitoring of a Hybrid Energy Storage Device	GM Global Technology Operations LLC
JP2012043682A (2010)	Battery Pack System	Toshiba Corp.

+ 10 more documents from the same query, each linked to Espacenet

Source: DEEPTech-Novac, HBSC t-SNE Analysis + Cluster Explorer — Cluster 3 detail, 6 of 16 families shown

Why It Matters: WIPO Says So Too

«Clustering is a statistical technique that groups a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.»

A. Trippe for WIPO, Guidelines for Preparing Patent Landscape Reports, 2015, §6.5, p. 58

The Contribution of Embedded Artificial Intelligence

- ▶ Assisted classification of ambiguous documents
- ▶ Automatic summarisation of abstracts and claims
- ▶ Assistance in building SQL/Python queries
- ▶ It assists the analyst, it does not replace them: human judgement remains central

Conclusion

- ▶ Patstat/SQL already provides the quantitative backbone of a credible PLR
- ▶ TIP does not replace it: it extends it, in a single session, beyond the structural limits of SQL alone
- ▶ The result is a continuous workflow: from query to interactive dashboard
- ▶ Working faster is not just convenience: it shifts analyst time from data handling to search-strategy refinement

References

- ▶ A. Trippe for WIPO, Guidelines for Preparing Patent Landscape Reports, WIPO, 2015 — wipo.int/edocs/pubdocs/en/wipo_pub_946.pdf
- ▶ WIPO, Patent Landscape Reports (publication series) — wipo.int/publications/en/series/index.jsp?id=137
- ▶ EPO, Technology Intelligence Platform (TIP) — tip.epo.org
- ▶ Chart examples: DEEPTech-Russo, DEEPTech-Novac, DEEPTech-FlowNano project folders

Thank You

Questions?



APPENDIX

TIP & Artificial Intelligence for Patlib Centres

SWOT: Making TIP Accessible Through AI

A SWOT snapshot from the EPO TIP Working Group — bridging non-programmers, patent data and generative AI (July 2026)

Strengths

- ▶ Domain experts can use plain-language questions — AI handles SQL/Python
- ▶ No-code tools (e.g. “PATSTAT Explorer”) built in days, not months
- ▶ Structured guides correct AI’s recurring, predictable mistakes
- ▶ Shows the reasoning path, not just the finished output — builds trust

Weaknesses

- ▶ AI is confidently wrong — errors are hard to catch without deep data knowledge
- ▶ Copy-pasted, unverified AI queries risk client-facing data quality
- ▶ Time-poor participants still struggle to adopt, even with AI support
- ▶ External data integration (SME size, precise geolocation) remains unsolved

Opportunities

- ▶ AI-generated lead lists and regional analyses for Patlib prospecting
- ▶ Reusable guides scale training without one-to-one teaching
- ▶ Custom AI models on EP full text as a further literacy step
- ▶ The “bridge” model is transferable well beyond patents and Patlib

Threats / Caveats

- ▶ Over-reliance on AI without human review can erode client trust
- ▶ External AI tools transmit data to third parties — never use with confidential data
- ▶ Adoption gap may widen between power users and casual participants
- ▶ Expert time is finite; material must be maintained as TIP evolves