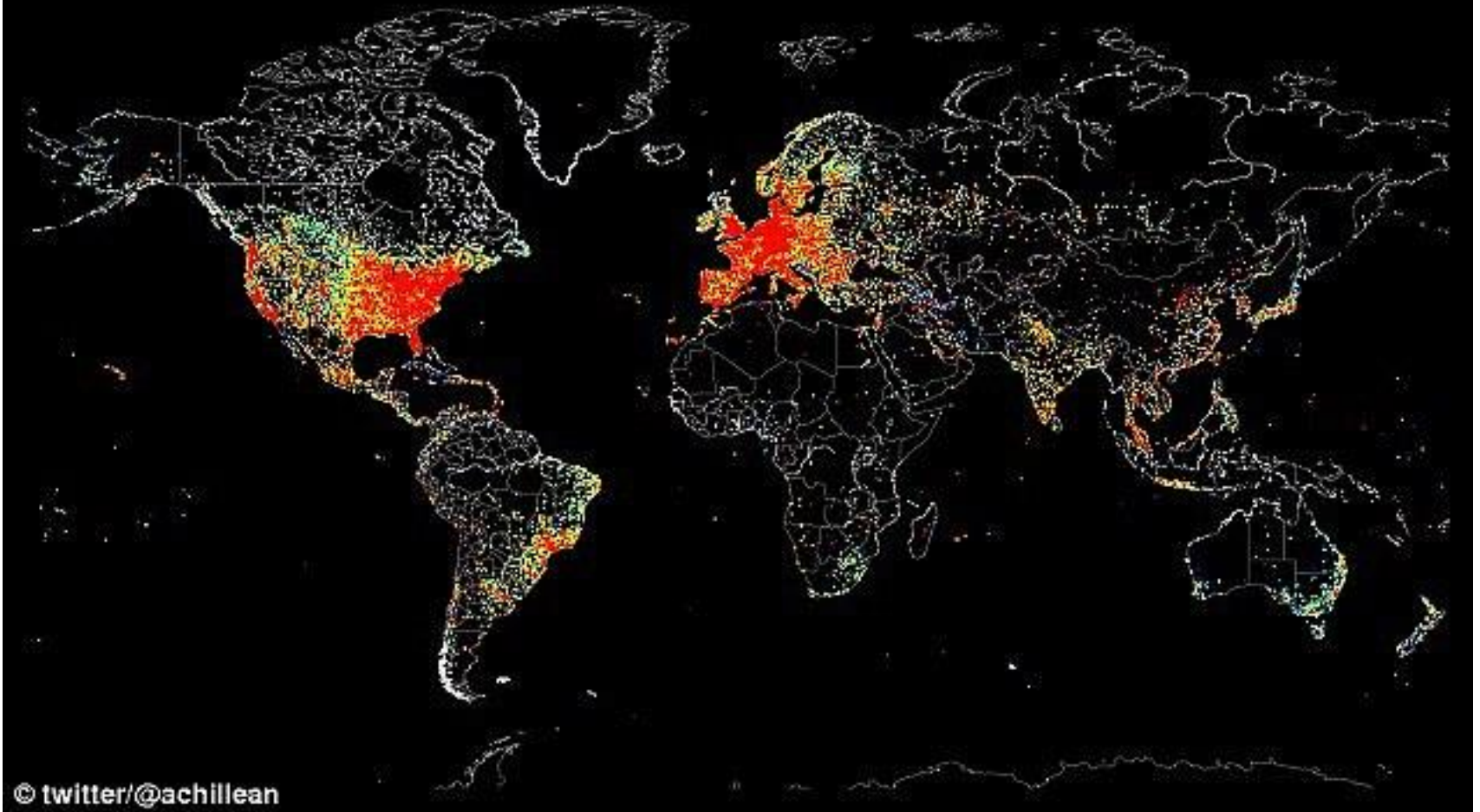


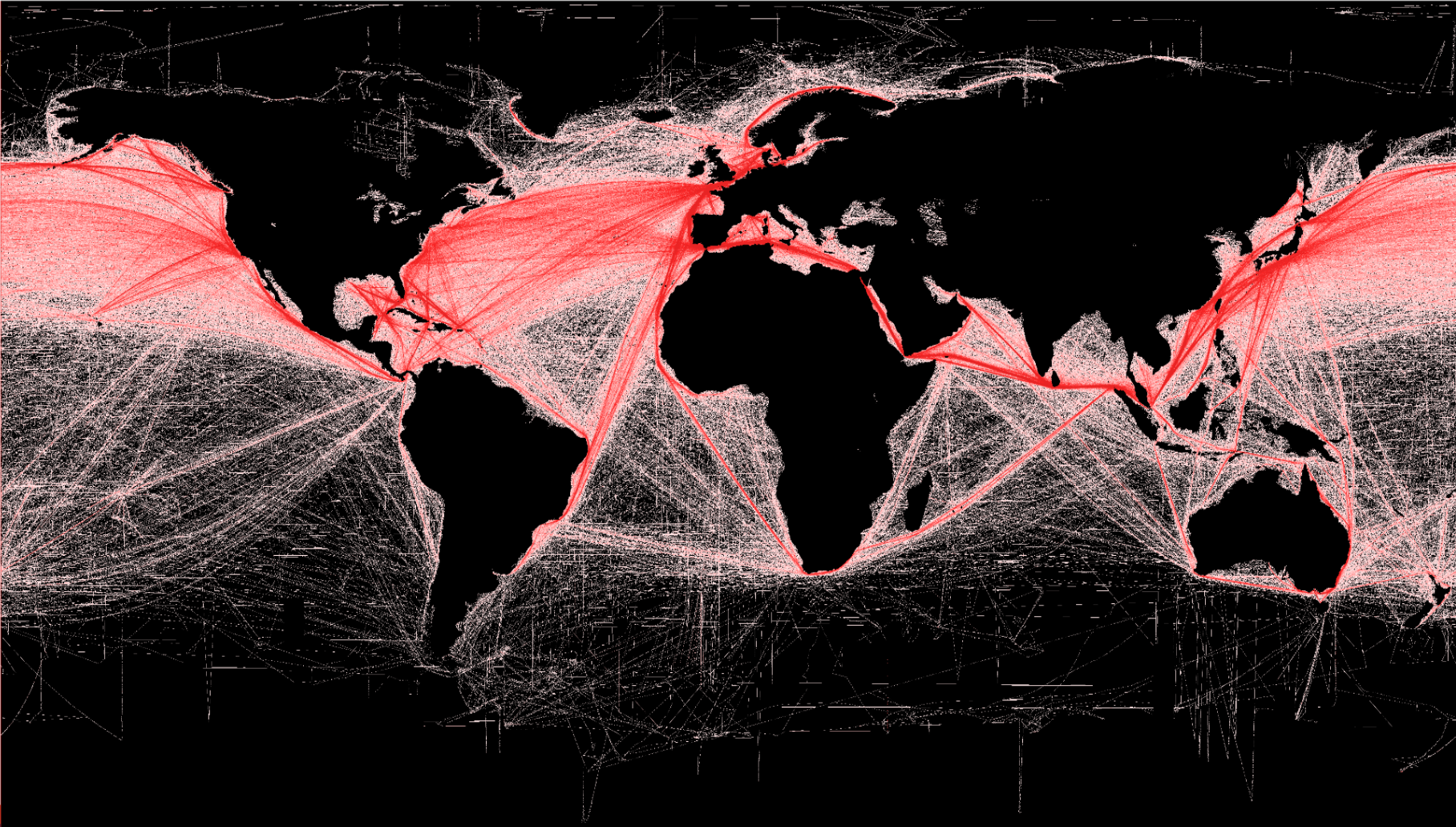
Text and Data Mining: How Visualizations Facilitate New Scholarship.

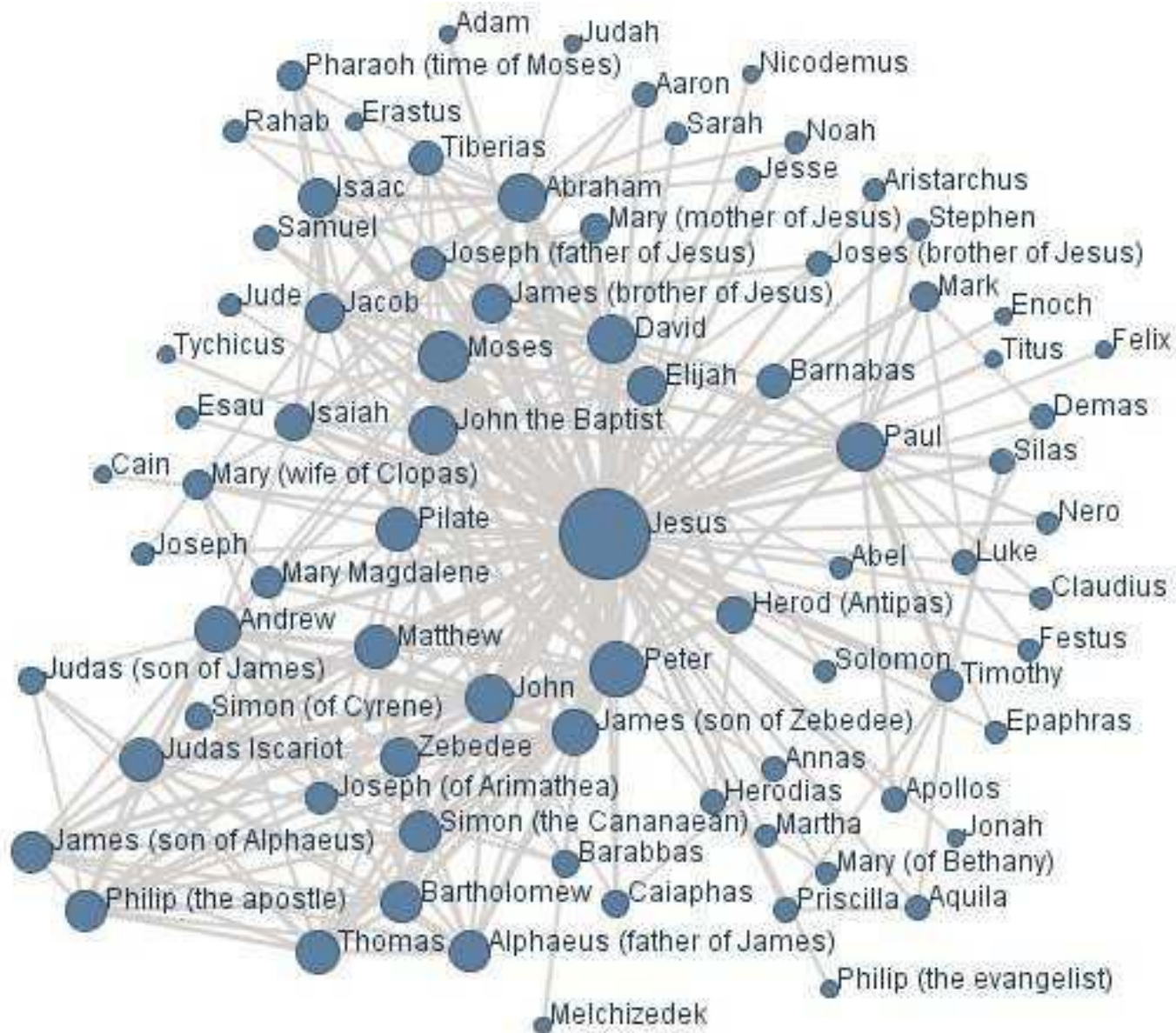
Gary Pollack
EBSCO Information Services

SHODAN



© twitter/@achilleian





<https://world.twingly.com/>

- **Introduction: EBSCO's approach to Text & Data Mining (TDM)**
- **Conversations with Publishers and Institutional Partners**
 - Spreading awareness of collections, metadata, and capabilities
 - Ensuring data security and ownership
- **Examples of current scholarship with EBSCO's Historical Digital Archives and TDM**
 - working within rights considerations
 - many different kinds of analyses possible with collections and new technologies
- **TDM scholarship and EBSCO's contemporary collections**
 - in development many potential audiences to be reached, including: professionals, educators, and students
 - no tech and low tech without leaving EIS environment
- **TDM Initiatives from EBSCO**
- **Wrap-up and questions**

How is EBSCO defining Text and Data Mining?

Using technology to identify, extract, and assess useful information in the form of patterns, trends, or relationships from large datasets.

These datasets can include collections of unstructured text and structured data. Results are used in the educational, scientific, and professional realms for illustration, explanation, example, comment, criticism, teaching, research, and analysis.

Why is EBSCO getting involved with Text and Data Mining Projects?

- To be a part of new knowledge sharing among Scholars, Researchers, and Digital Humanists
- To meet increasing demand of Scholars wanting to mine the larger corpus of EBSCO datasets
- To support this type of research in the undergraduate curriculum and prepare students with the analytical skills that they will need in the future.
- To meet growing demand of libraries and publishers
- To play an active role in establishing industry models, standards, and best practices
- To leverage robust and valuable datasets that already exist for Digital Archives, Magazine Archives, and contemporary journals and magazines

What requests have EBSCO's Customers made for TDM?

- Customers have submitted various requests for TDM: for example:
 - Text & Data mining request from Harvard scholar via Librarian for access to text and data of 2 periodicals from AAS Historical Periodicals Collection
 - Purdue Professor's request to mine the Illustration metadata from the AAS collections using a team of undergraduate interns
 - UConn scholar via AAS request for metadata for subject terms and full text for title *Godey's Lady's Book* from AAS Historical Periodicals Collection for book research
 - Ole Miss scholar request for all text and data from Business Source Premier database to do Sentiment Analysis
 - German scholar request for download of all data/text from 3 contemporary business journals for last 15 years to study high quality public management research
 - MIT wants to know if EBSCO has tools that allow “plumbing” data/text from EBSCO collections for their scholars
 - Stanford wants to “harvest” all articles published and cited in MLA by their scholars and faculty

Why EBSCO Is Well-Suited for TDM?

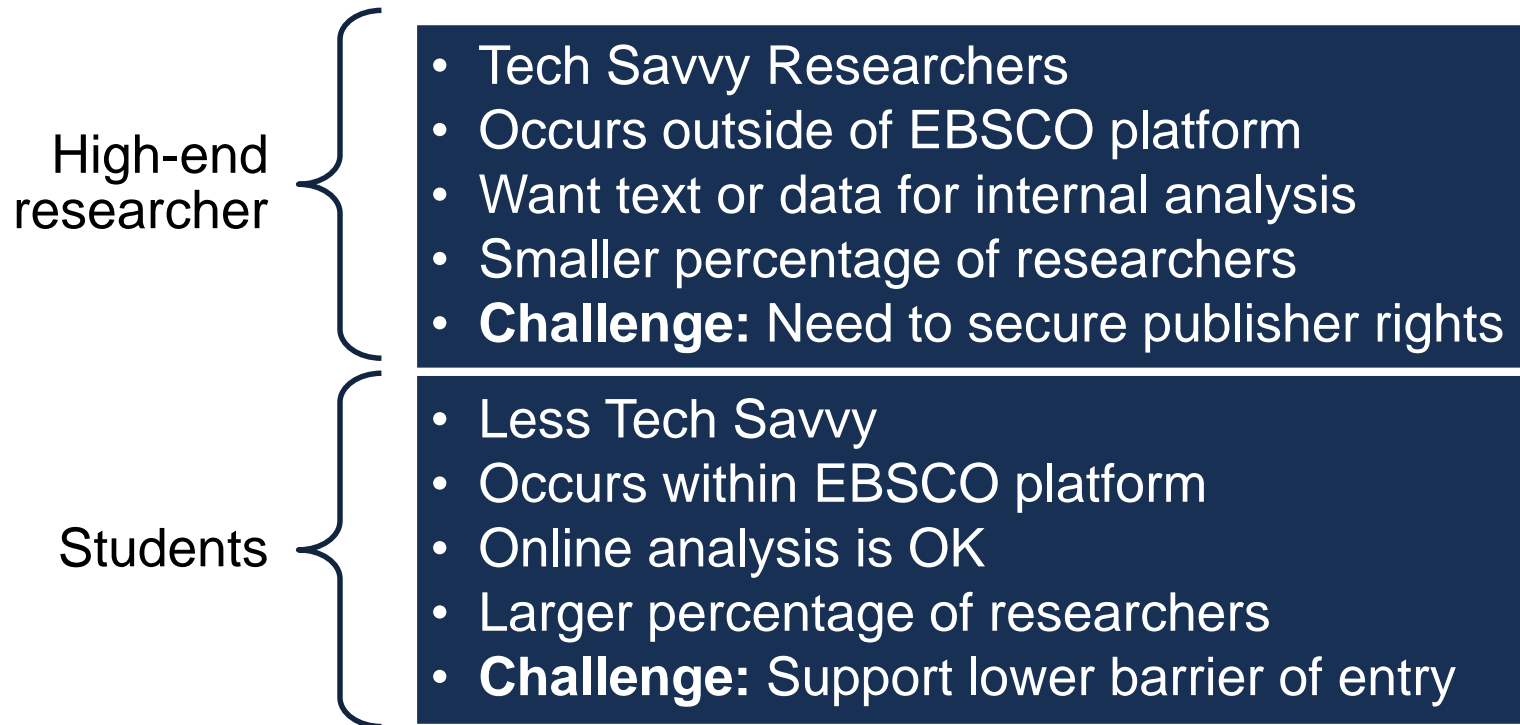
- Rich metadata, depth and breadth of content
- Strong relationships with Publisher partners
- Rights management of Full Text
- Technology (API, interfaces, indexing, search, product integration)
- Understanding the **Academic Audience** (Contextual Inquiry, UX, Pilots, Advisory)

Why are TDM skills important?

- EBSCO is identifying needs and research skills of different kinds of users
 - undergraduate students, as part of building new research skills
 - Pedagogical changes in the undergraduate curriculum
 - Teaching and learning analysis skills in a changing world
- Results for different users:
 - Academia: new scholarship and insight during the research experience
 - Corporate: trend analysis and market evaluation
 - Schools: critical thinking and expanded research options

What are the Challenges?

- Supporting two types of TDM researchers



- No clear-cut standardized models or best practices

Challenges continued...

- Rights Management
 - Dialog with publisher partners - Trust
 - Honoring publisher restrictions
 - Secure sharing of data
 - Terms of use with partnering scholars
- Tools for researchers and students
 - Search against Products (Full Text and Metadata)
 - Intuitive
 - Discoverable
 - Visualized results

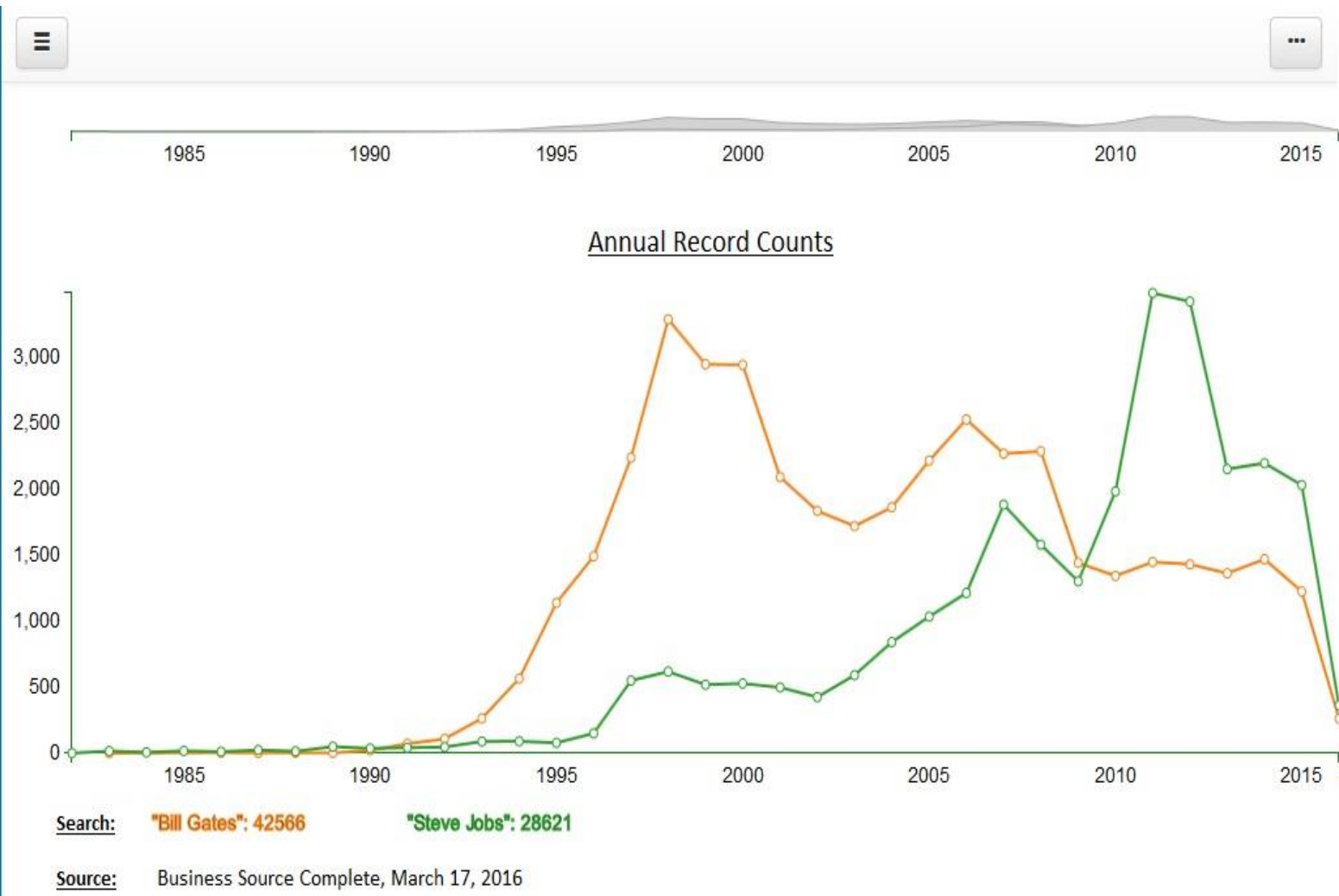
TDM: Text & Data Mining

"Bill Gates" +

"Steve Jobs" x

More options

Search



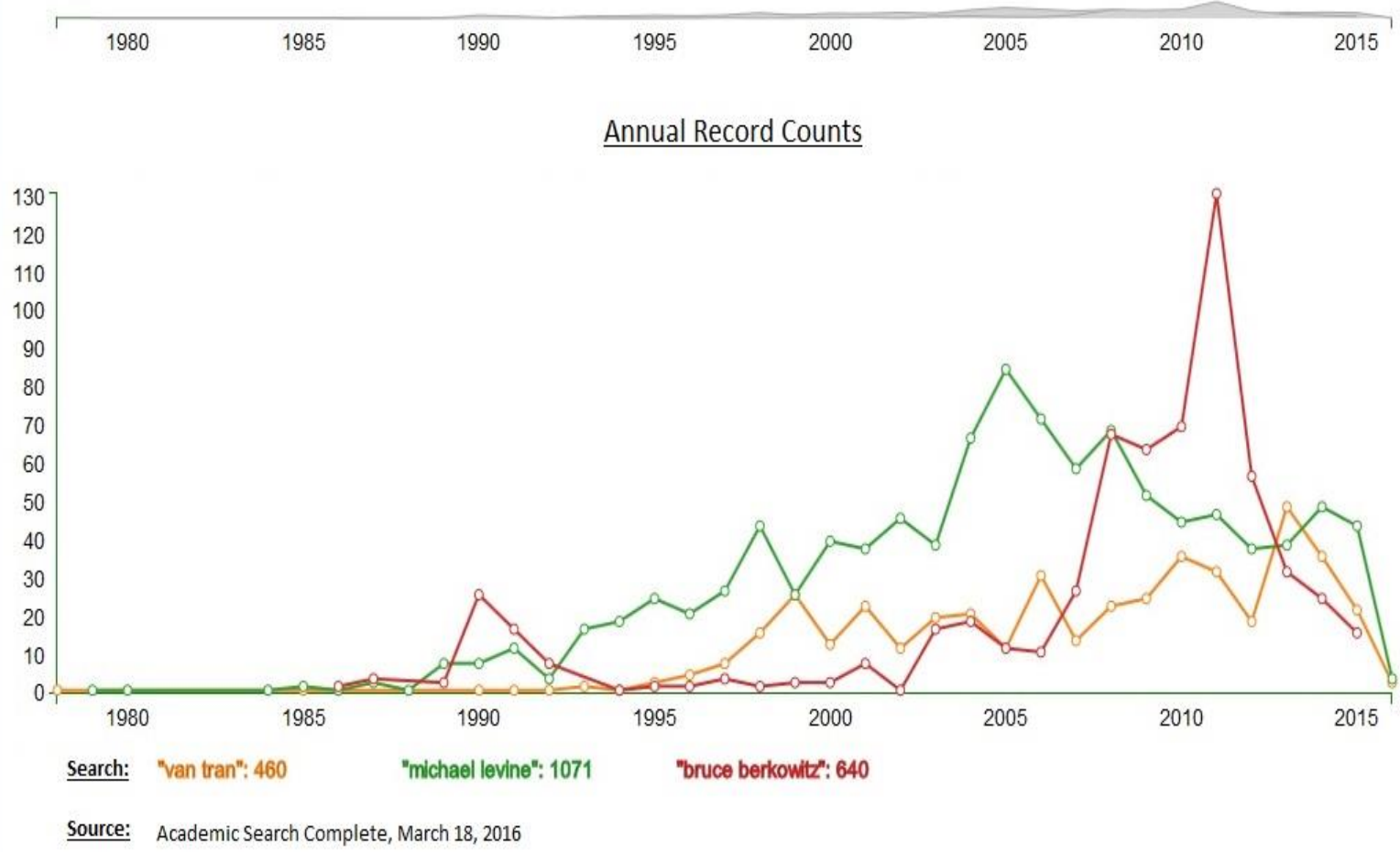
TDM: Text & Data Mining

"van tran" 669 +

"michael levine" 669 x

"bruce berkowitz" 669 x

More options Search



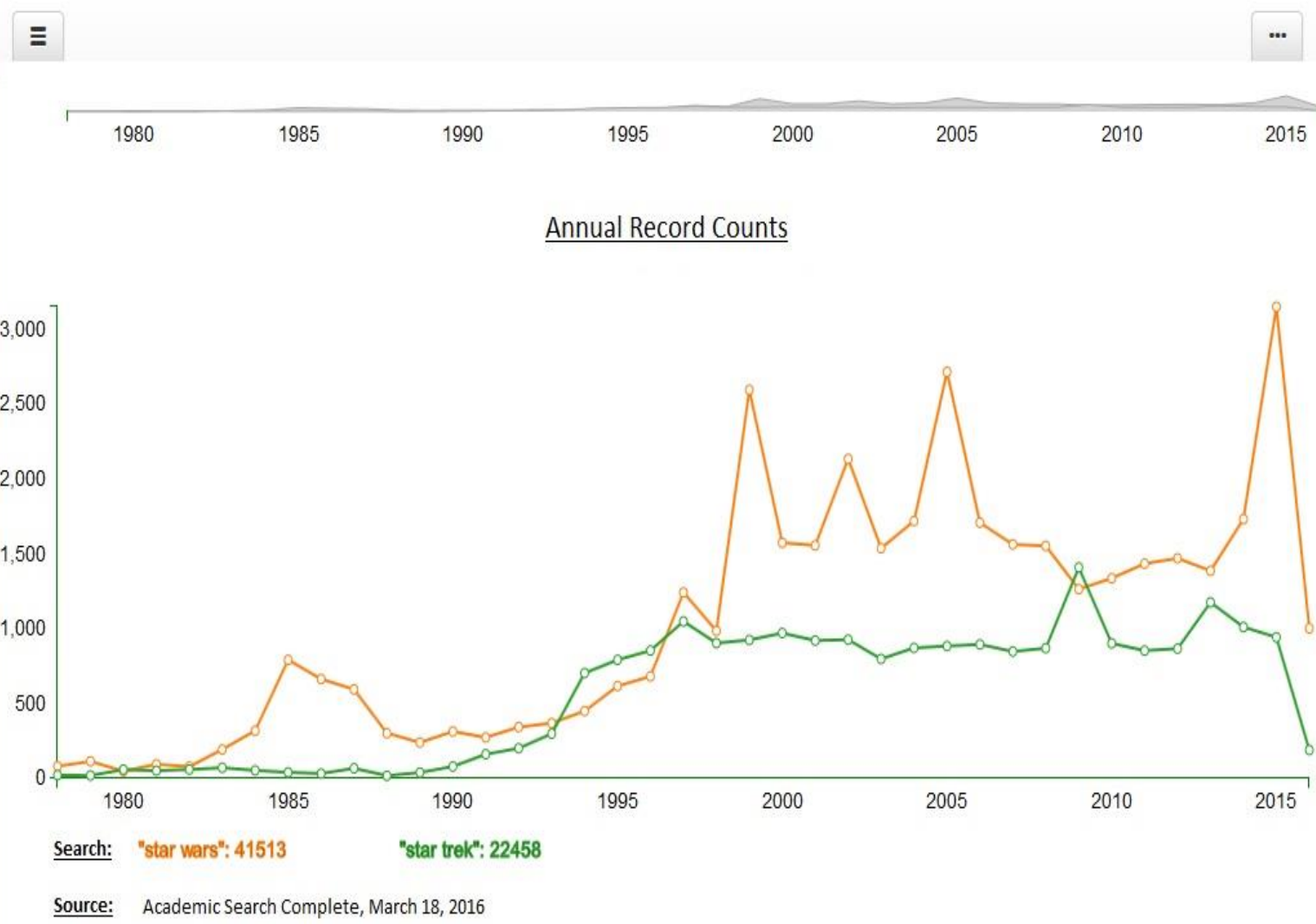
TDM: Text & Data Mining

"star wars" [icon] +

"star trek" [icon] x

More options

Search



EBSCO's Goals and Objectives

- Release initial Data Visualization product
 - Trend analysis over time visualization
 - Beta Pilot with small group of customers
 - For small set of products
 - Feedback to position for product roll-out
- Provide data delivery services
 - Perpetuity Digital and Magazine Archives
 - Terms of Use sign-off
- Establish framework/foundation for longer-term plan/solutions

Questions?

Email: ETDMServices@ebSCO.com

Thank You!

Gary Pollack
EBSCO Information Services