Turning Visions Into Reality:
Lessons Learned From Building a Search and Discovery Platform

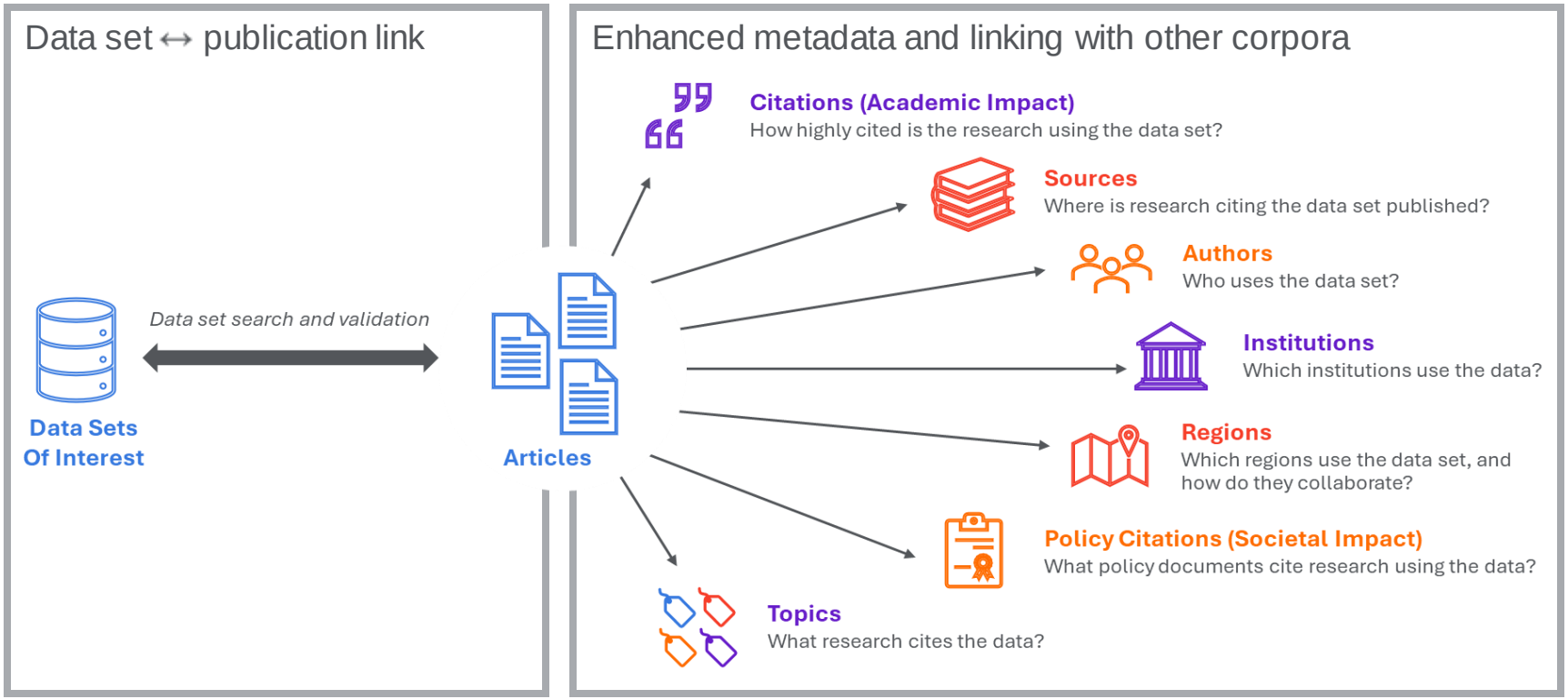
ELSEVIER

JOHNS HOPKINS
INSTITUTE for
DATA-INTENSIVE
ENGINEERING & SCIENCE

TACC
TEXAS ADVANCED COMPUTING CENTER

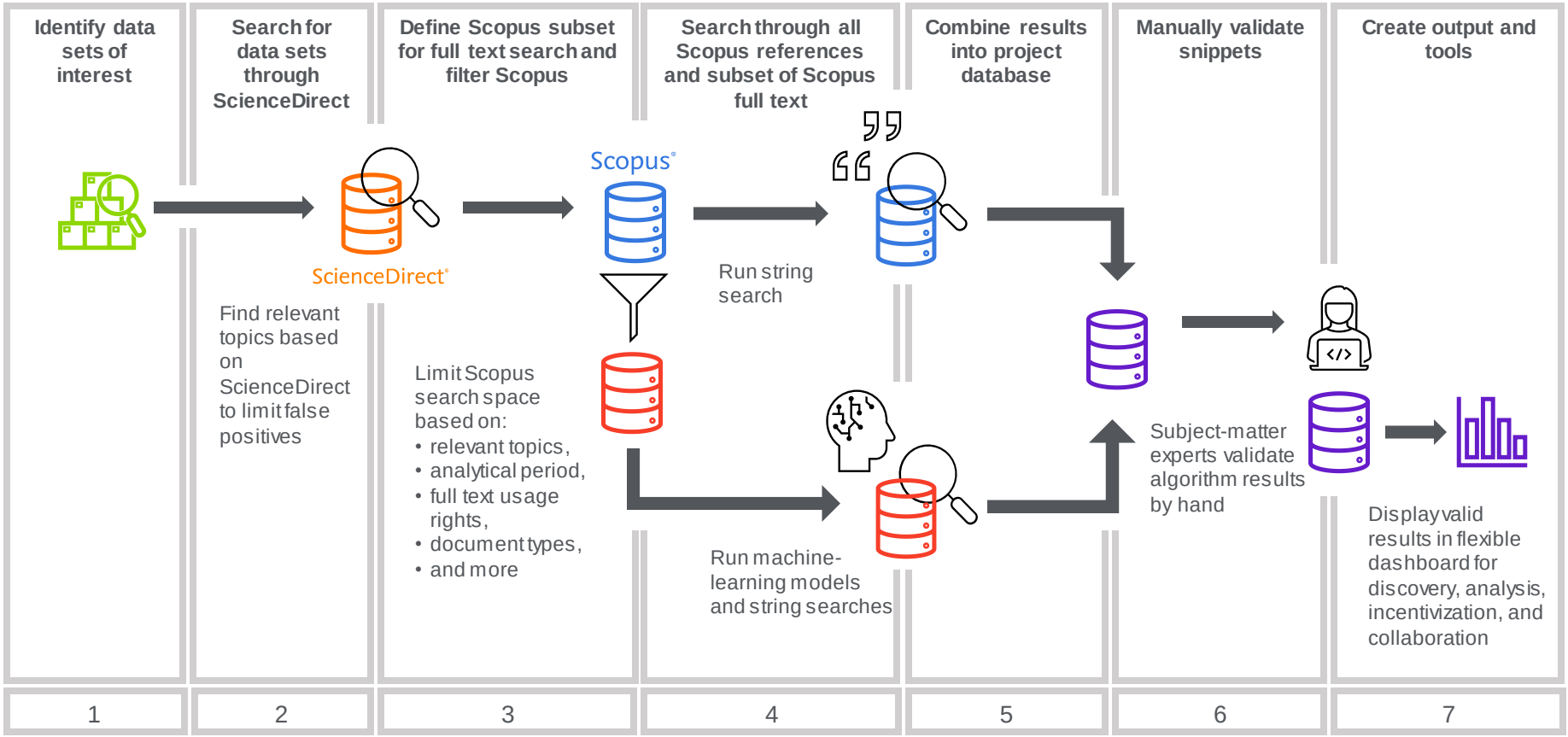
democratizing data

Context: Search & Discovery Model



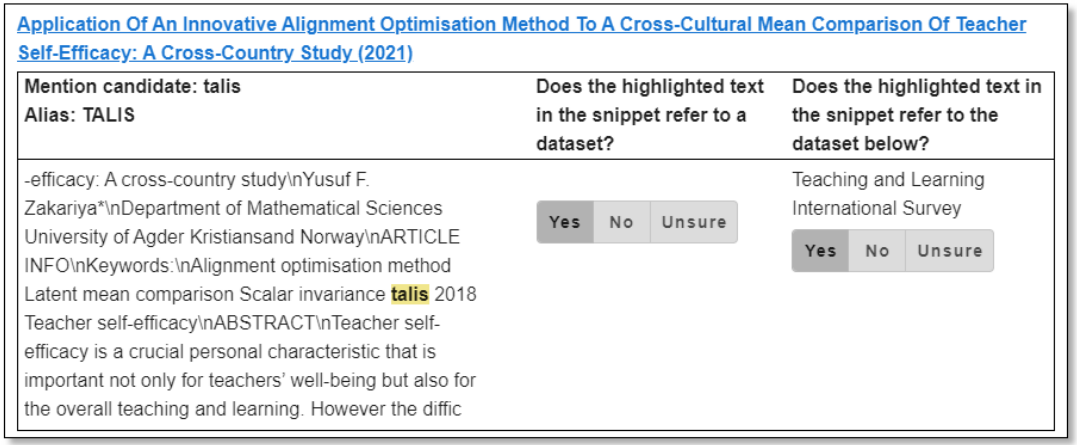
Articles: 96 million+ records in Scopus

Process Description: Searching for Datasets



Validation: Text Review by Subject Matter Experts

Snippet Validation Example



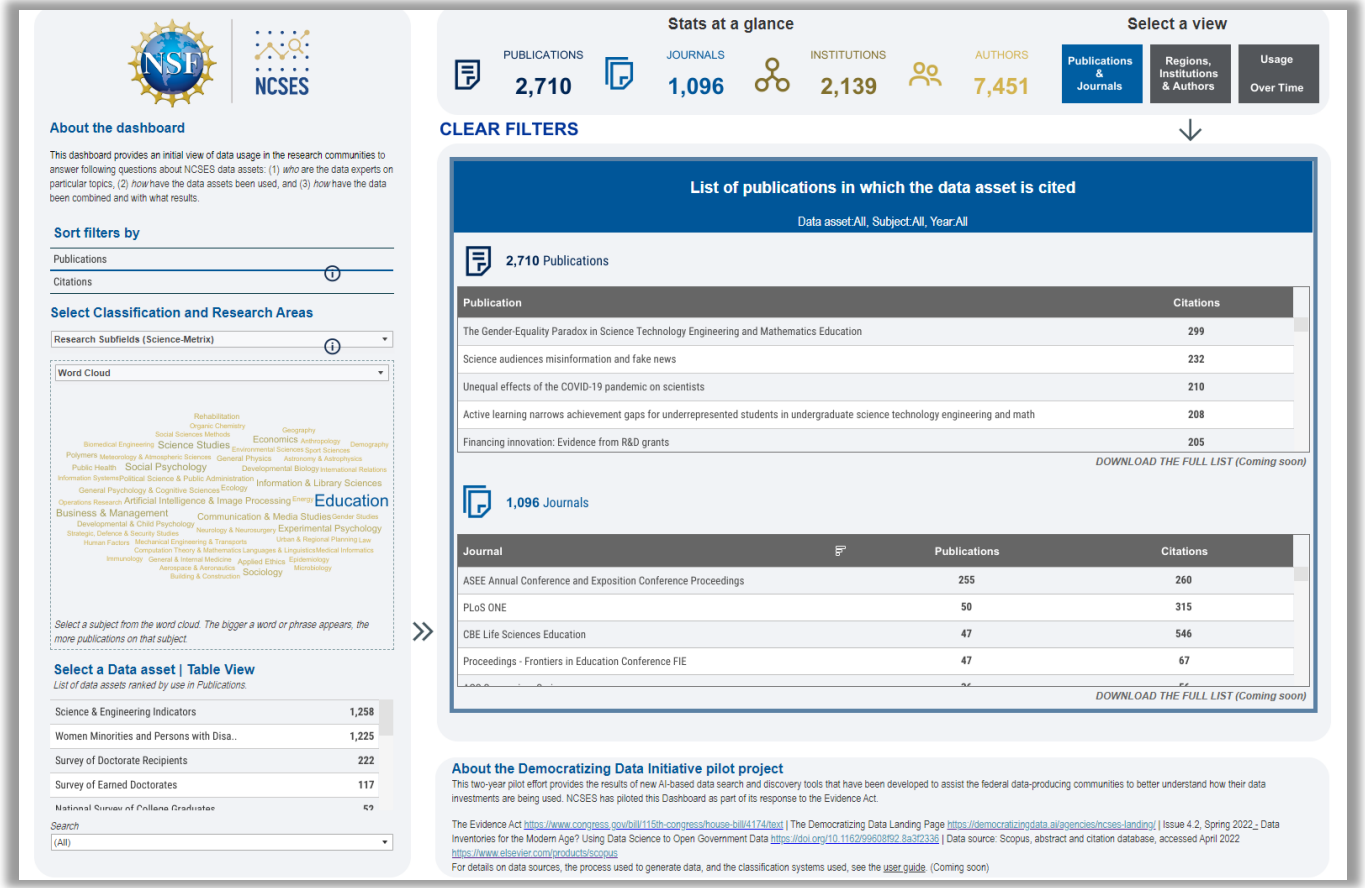
Application Of An Innovative Alignment Optimisation Method To A Cross-Cultural Mean Comparison Of Teacher Self-Efficacy: A Cross-Country Study (2021)

Mention candidate: talis Alias: TALIS	Does the highlighted text in the snippet refer to a dataset?	Does the highlighted text in the snippet refer to the dataset below?
-efficacy: A cross-country study\nYusuf F. Zakariya*\nDepartment of Mathematical Sciences University of Agder Kristiansand Norway\nARTICLE INFO\nKeywords:\nAlignment optimisation method Latent mean comparison Scalar invariance **talis** 2018 Teacher self-efficacy\nABSTRACT\nTeacher self-efficacy is a crucial personal characteristic that is important not only for teachers' well-being but also for the overall teaching and learning. However the diffic	Yes No Unsure	Teaching and Learning International Survey Yes No Unsure

Validation Results Example

Data Asset / Alias name	Actively Validated	Positively Validated	Precision
Higher Education Research and Development Survey	42	42	100%
HERD	1308	8	<1%

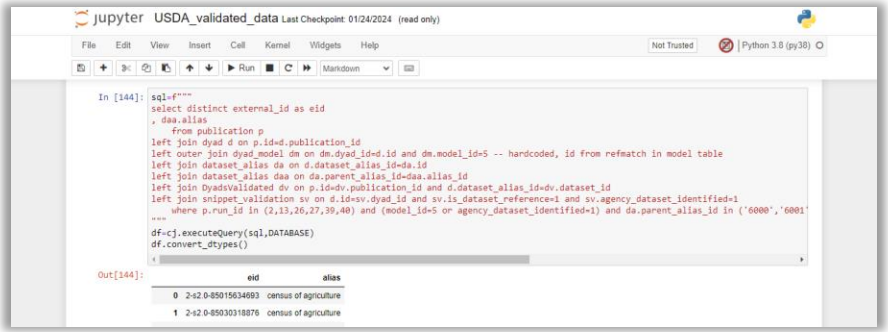
Access & Dissemination: User-Friendly Dashboards

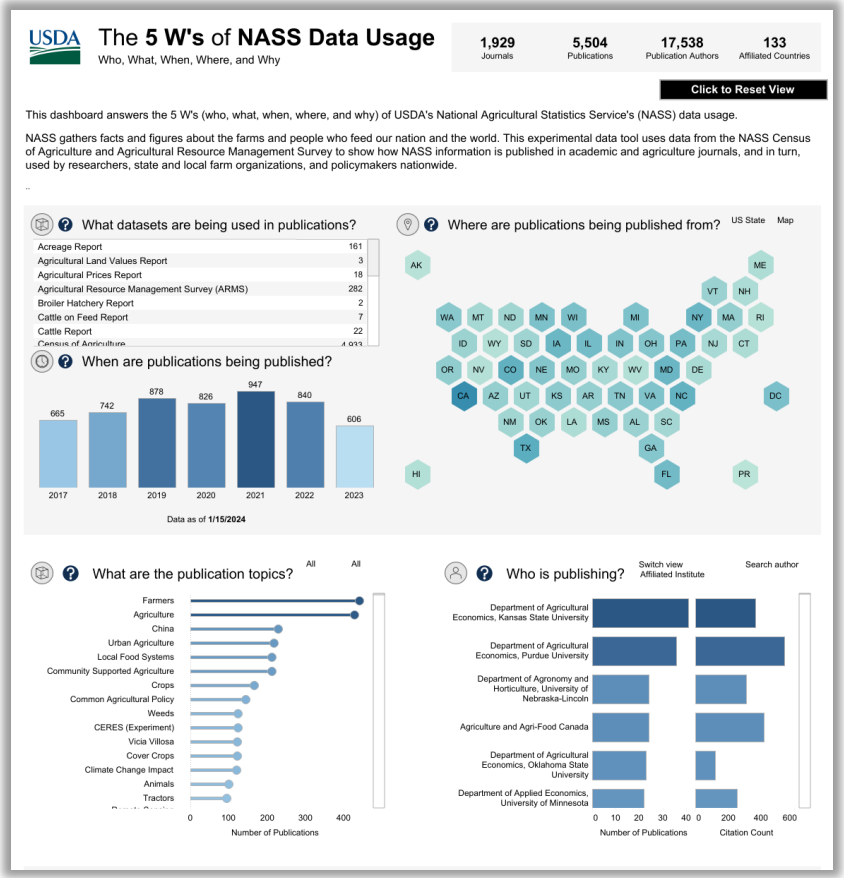


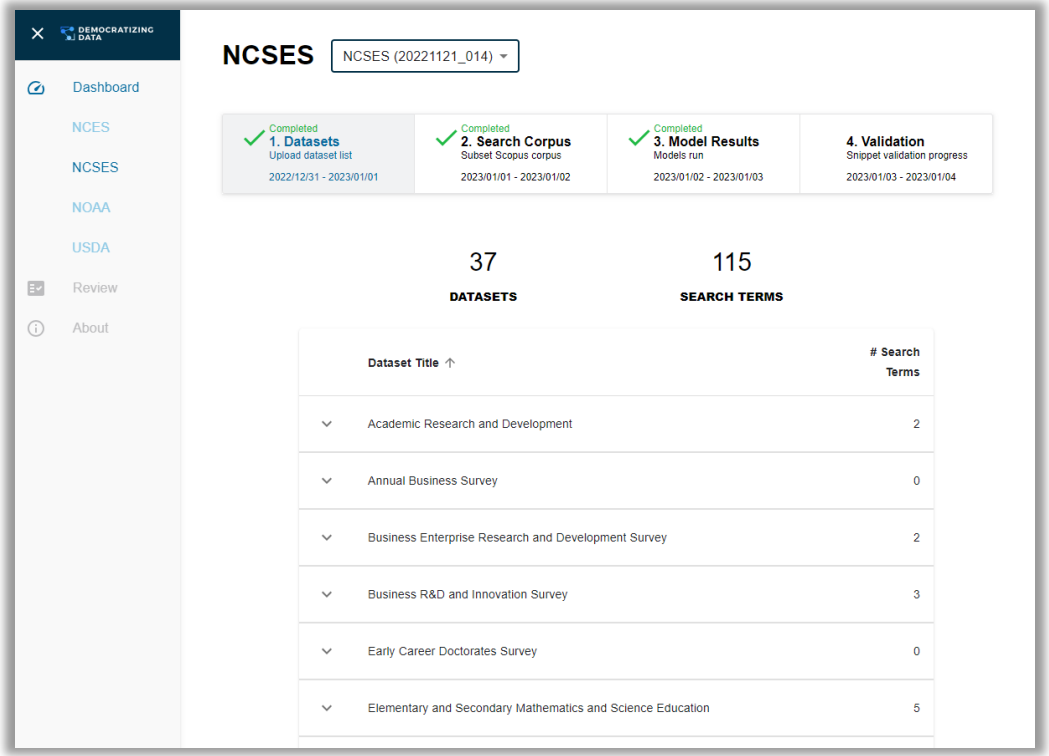
https://democratizingdata.ai/

Access & Dissemination: Tools for Custom Analysis

- SciServer (cyberinfrastructure system with tools and big data)
- Application Programming Interface (API)





Future Directions

- Increase transparency and collaboration through administrative dashboard
- Expand to include other corpora (e.g., policy documents)
- Develop meaningful, tailored classifications through community engagement
- Improve data usage measures
- Refine user tools