Profile of States of USA: Nanotechnology Research Emergence

Alan Porter, Mark Markley, Nils Newman & Ric Snead
Search Technology, Inc.

aporter@searchtech.com
404-384-6295

We acknowledge US National Science Foundation support through Award #759960 – "Indicators of Technological Emergence" to Search Technology, Inc., and Georgia Tech. The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Profiling Cutting-Edge Research: The Case of Nanotechnology by States

1. Introduction to Tech Emergence Indicators
2. Gauging States' cutting edge nanotechnology research

Part 1. The Tech Emergence Indicators (TEI)

- **Aim --** practical indicators that discern ***cutting edge*** R&D development, ***within a tech domain***
  - Could present together with other R&D measures that distinguish frontier R&D activity
- **Heritage**
  - **IARPA FUSE** (U.S. Intelligence Advanced Research Projects Activity, Foresight & Understanding from Scientific Exposition) Program promoted emergence indicators development (2010~14)
  - **NSF Science of Science: Discovery, Communication, and Impact (SoS:DCI),** formerly **SciSIP,** 3 awards to Search Technology and Georgia Tech Program in Science, Technology & Innovation Policy(STIP)
  - Current TEI project also supported by **NSF National Center for Science and Engineering Statistics**

TEI Criteria

Drawing on IARPA FUSE &
Rotolo et al., What is an emerging technology? (2015)

- 4+1 attributes as **thresholds**:
  - Term Novelty
  - Term Persistence
  - Term Growth
  - Research Community
  - *+Scope (specificity to technology field of interest)*

- **Accelerating Trends** -- Primary Tier: Calculation of "EScores" for each candidate **term**
  - Our formulation – 3 base + 7 active = 10 periods (e.g., years) of data
  - $\sum((2 * Active),\ Recent,\ Slope)$

TEI: 2 Tiers

- **Primary Tier –**
  Distinguish **terms** (and/or consolidate into **topics)** that evidence **recent, accelerating R&D attention**

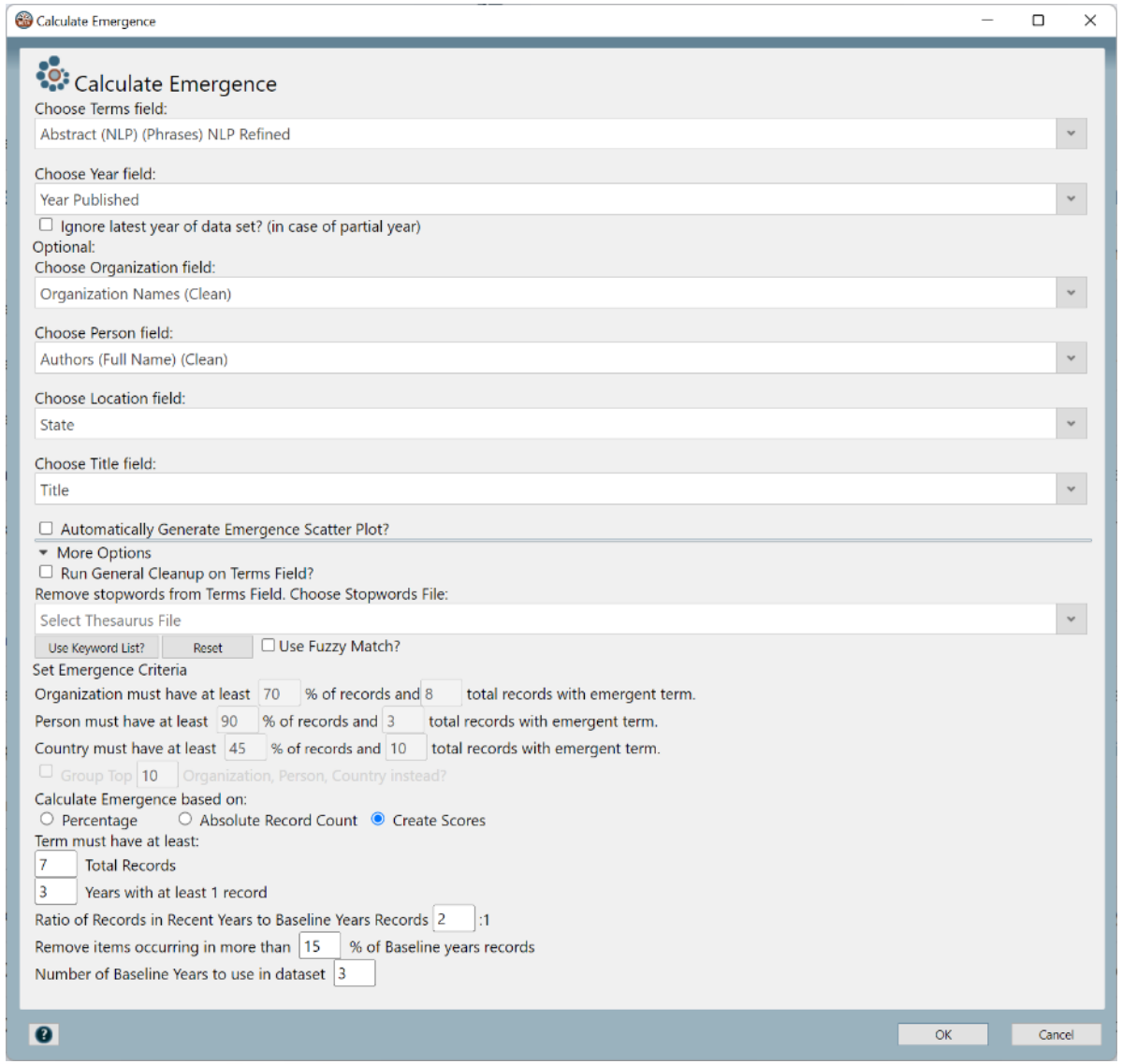
- **Secondary Tier –**
  Aggregate emergent terms per abstract records to distinguish **players most actively researching such cutting-edge topics**
  [“players” can be countries, organizations, or authors – or, for NCSES, U.S. states]

How We Generate Tech Emergence Indicators (TEI)

Distinguish topics growing rapidly in usage in a target domain - [U.S. Patent 10,803,124]

1. Specify the Science or Technology domain & suitable database source (e.g., Web of Science)
2. Search & retrieve abstract records
3. Select topical fields (e.g., Title & Abstract Natural Language Processing phrases)
4. Refine terms
5. Generate emergence scores (**EScores**) for terms
6. Generate EScores for players, based on use of those terms in their authored paper titles & abstracts (e.g., **states**)

Calculating Emergence Scores
(Porter et al., *TF&SC*, 2018; Carley et al., *Scientometrics*, 2018)



Calculating Players' TEI Scores (Secondary Tier)

- Total Emergence Indicator =
  Sum [SQRT (Emergent term score) X # of records authored by that player containing the term]
  ** only count terms scoring >1.77
  ** reflects overall R&D effort addressing cutting edge terms (topics) in the domain by that player

- [optional alternative] Normalized Emergence Indicator =
  Total divided by the SQRT (# of records by that player)
  ** gets at intensity of cutting edge R&D effort in domain [i.e., relative emphasis on frontier terms (topics)]

- [See Porter et al. (2018) and Carley et al. (2018) – in References]

Our Nanotechnology Data

- Using Web of Science Core Collection

  [standard web access, although Georgia Tech has now licensed full XML access]

- Complex Boolean search refined over the years – most recently:

  Wang, Z., Porter, A.L., Kwon, S., Youtie, J., Shapira, P., Carley, S.F., and Liu, X. (2019), Updating a search strategy to track emerging nanotechnologies, *Journal of Nanoparticle Research* 21: article #199; doi.org/10.1007/s11051-019-4627-x.

- Yielding 2.2 million records for Years 1991-2019

Step 1: Refine the Nano Data

- Extract the U.S. authored (or co-authored) nano records through 2017

  [our more recent data are not consistent without extensive rework]
- Use 1999-2017 for the profile → 404,706 records
  - 99.97% containing abstracts
- Use *VantagePoint* text analysis software to generate Natural Language Processing (NLP) noun phrases from the abstracts to get at topics addressed
  - Alternatively, one could combine with title NLP phrases and (usually not) keywords

Step 2: Refine the Nano Data

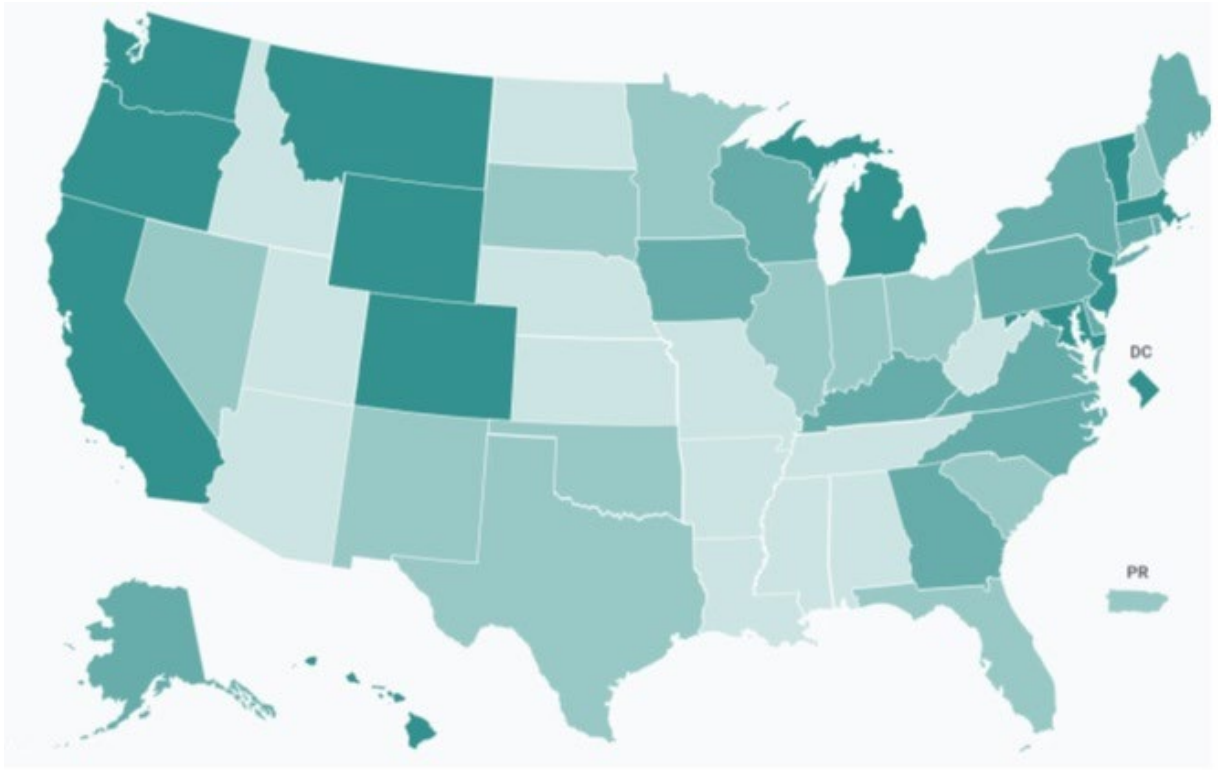
- Run *VantagePoint's* "RefineNLP" process to refine terms
  - Optionally remove or keep words from user's selected keywords file
  - Remove spurious XML encoding
  - Simplify Chemical Compounds
  - Remove general stopwords
  - Remove common scientific and academic stopwords
  - Run conservative stemming
- Generate suitable fields for US authors, US author organizations, author states
- Consolidate name variations for authors and author affiliations using fuzzy matching with rulesets
- Create sub-datasets for each of the 10-year periods – calculation of Escore for a given year requires 10 periods of data

Calculate the Nano TEI's

- Run the Escoring process (in *VantagePoint*)
  - Select a 10-year period
- Export Emergent Terms & scores to Excel
- Export State Emergence scores to Excel
  - As Data Table for NCSES 'sandbox' to generate desired outputs
  - Emergent terms with EScores available if want to profile domain (like the military technology study)
  - State Emergence scores available for each time period

States' Nanotechnology Research

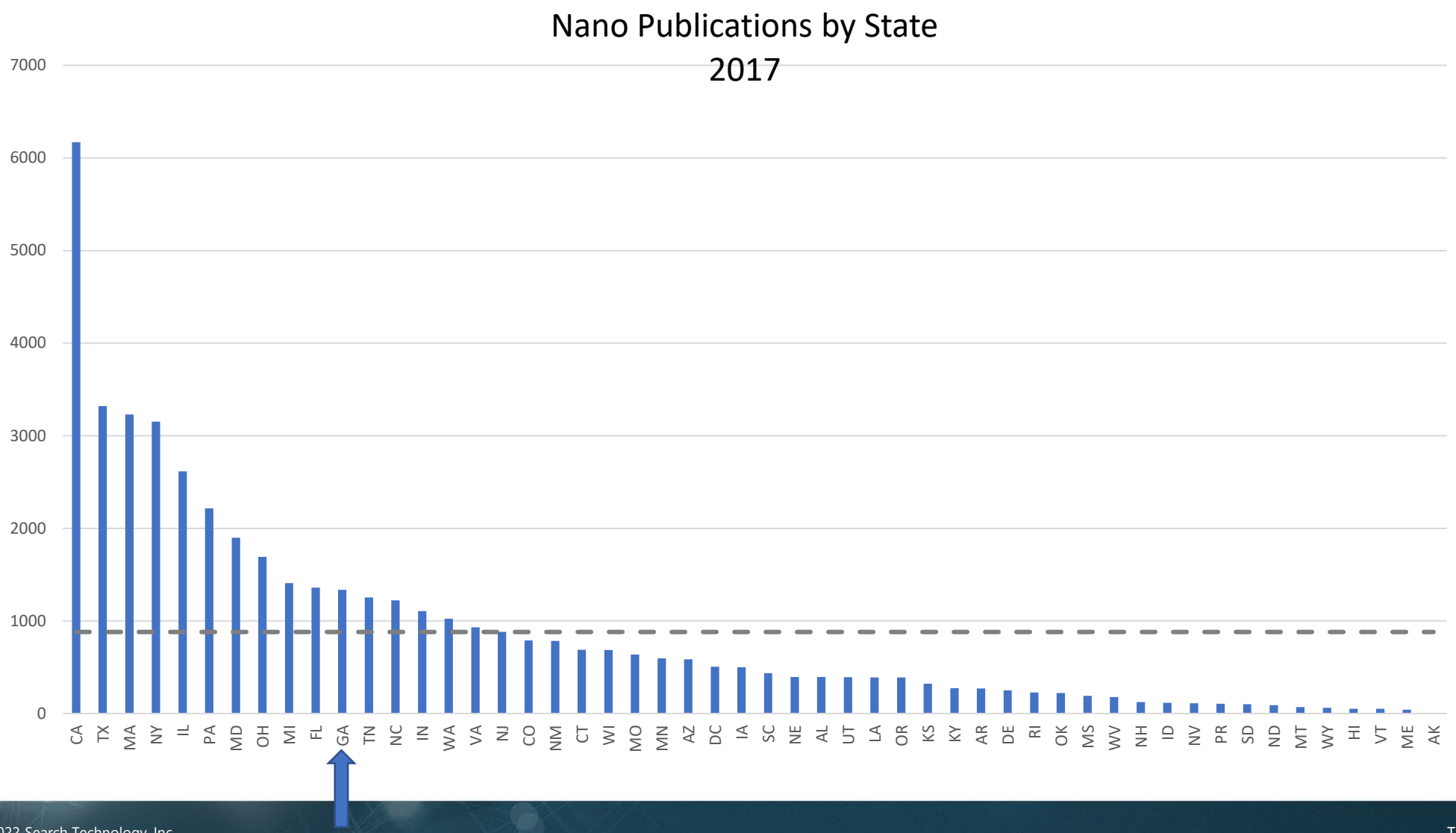
- Extract and refine 404,706 U.S. nanotechnology research publications from Web of Science
- Profile States' nanotechnology research
- Distinguish their Cutting-Edge (i.e., Emergent) nanotechnology research publication
- [Future] – Compare with other data by State – e.g., National Nanotechnology Initiative Funding



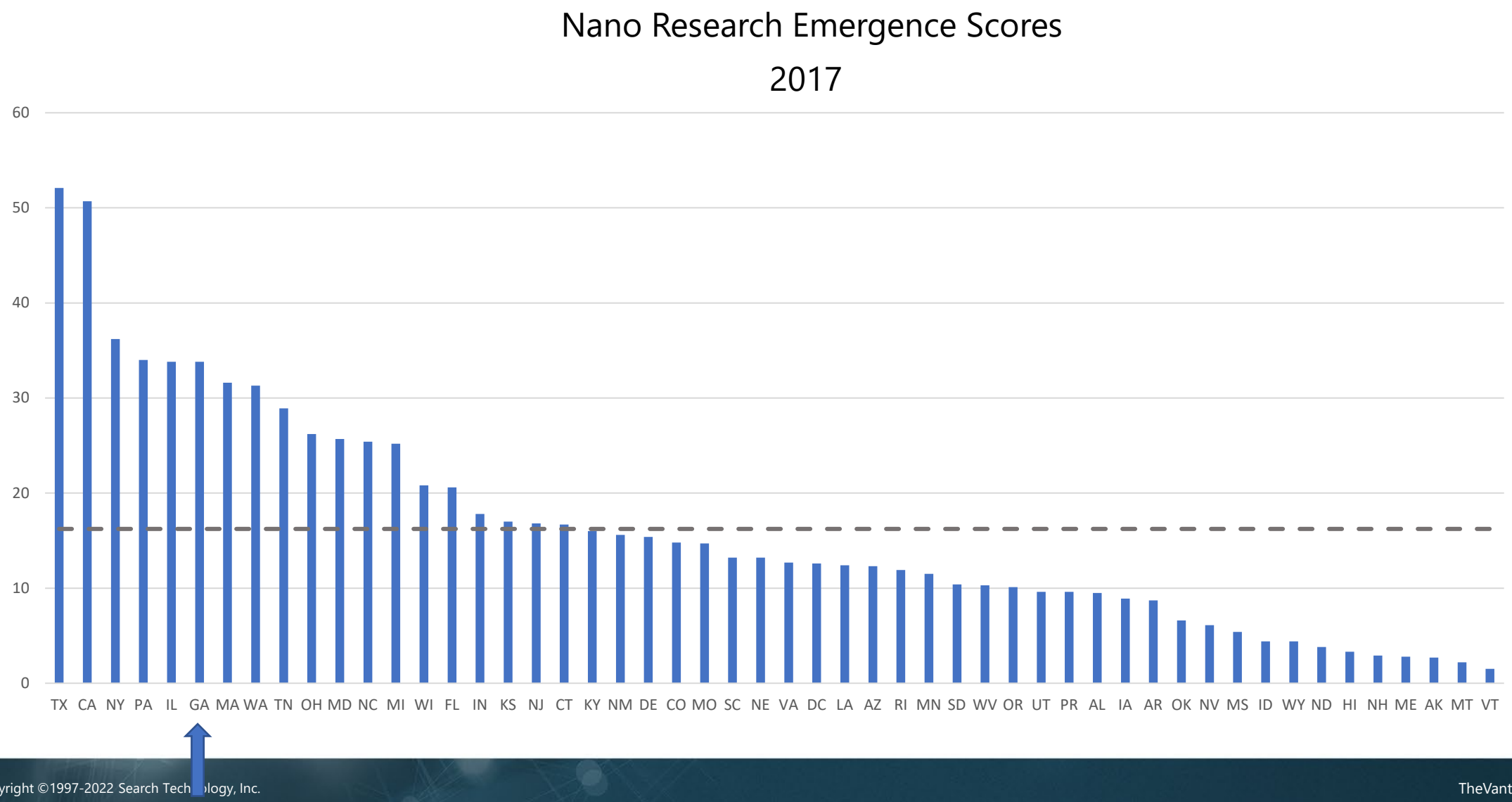
Selected States' Nano Research Papers & Emergence

2015			2016			2017		
Records	State	Emergence	Records	State	Emergence	Records	State	Emergence
21075	TX	78.2	23041	TX	80.7	24887	TX	52.1
44283	CA	73	46899	CA	73.8	49290	CA	50.7
22532	NY	64.4	24071	NY	64.9	25411	NY	36.2
21151	MA	57.9	23031	MA	56.3	17607	PA	34
17241	IL	46.4	16646	PA	48.4	19603	IL	33.8
15578	PA	45.8	18393	IL	47.4	10108	GA	33.8
8733	GA	45.6	9531	GA	45.9	24704	MA	31.6

State Level Nano Publishing



State Level Nano Research Emergence Scores



Policy Applications: State Case

- TEI offers an important complementary metric.
- For nanotechnology funding policy, knowing State cutting-edge (emergent) R&D performance -- and rate of change -- complements knowing State nano R&D publication
- Other metrics would further enrich the bases for policy decisions – e.g., patent emergence scores; State funding through the National Nanotechnology Initiative
- Would be suitable to apply to many S&T domains – e.g., U.S. Critical & Emerging Technologies, such as hypersonics or quantum information technologies.
- Could inform choices re: fairness – to boost cutting-edge R&D to compete with foreign efforts. Or, boost weaker States.

TEI References [with cite counts in Web of Science as of 7/17/2022]

- Porter, A.L., Garner, J., Carley, S.F., and Newman, N.C. (2018), Emergence scoring to identify frontier R&D topics and key players, *Technological Forecasting and Social Change;* https://doi.org/10.1016/j.techfore.2018.04.016 [30 cites]
- Carley, S.F., Newman, N.C., Porter, A.L., and Garner, J. (2018). An indicator of technical emergence, *Scientometrics,* 115 (1), 35-49; http://link.springer.com/article/10.1007/s11192-018-2654-5 [33 cites]
- Kwon, S., Liu, X., Porter, A., & Youtie, J. (online, 2019), Research addressing emerging technological ideas has greater scientific Impact), *Research Policy*, 48 (9); https://www.sciencedirect.com/science/article/pii/S0048733319301544 [15 cites]
- Rotolo, D., Hicks, D., and Martin, B. (2015), What Is an Emerging Technology?" *Research Policy* 44: 1827–43; https://doi.org/10.1016/j.respol.2015.06.006. [293 cites]
- Li, M.N., Porter, A.L., and Suominen, A. (2018), Insights into relationships between disruptive technology/innovation and emerging technology: A bibliometric perspective, *Technological Forecasting and Social Change* 129, 285-296. [56 cites]
- Zhang, Y., Huang, Y., Chiavetta, D., and Porter, A. 2021, Tech mining for engineering management: An introduction, *IEEE Transactions on Engineering Management* 68 (5), 1211-1213; DOI: 10.1109/TEM.2021.3061862. [17 papers special section]
- Schmid, J. (2022), Technological emergence and military technology innovation, *Defence and Peace Economics*; DOI: 10.1080/10242694.2022.2076339.

Appendix

- Nano Data
- State level data possibilities

Data table (partial) for NCSES Template

State	2008	2009	2010	2011
Average	47.9	54.2	38.6	35.4
AK	10.3	14.1	11.8	6.4
AL	45.4	48.2	29.8	24.2
AR	38	47.8	38.4	35.1
AZ	40.4	47.7	33.8	29.1
CA	135.3	149.8	106.7	103.1
CO	32.9	42.1	28	25.4
CT	32.4	41.2	29.6	28.9
DC	47.3	47.4	32.1	31.1
DE	45.7	48.6	29.2	19.8
FL	72.1	79.5	56.1	52
GA	80.7	86.2	59.2	53.8

Part 3. Potential -- Explore NCSES Interests re: TEI

- TEI offers an indicator to measure players (e.g., states) engagement of cutting edge research within a target technology domain
- How might this enrich NCSES critical technology reports?
- Pilot?
  - Pick technology (e.g., Synthetic Biology)
  - Search (e.g., in WoS or Scopus)
  - Generate TEI outputs
  - Explore ways to incorporate in technology reports
- Additional ideas
  - Profile states across all Web of Science publications to determine primary emerging tech at the state level?
  - Profile citations received & emergence over time periods?