

DeSci Sensemaking Networks

Adding a sensemaking layer to DeSci

Ronen Tamari

DeSci ->	Broadening	access	to	science	

??? -> Making sense of all the new research DeSci is creating

The **X** factor in science

[Submitted on 15 Jun 2023 (v1), last revised 24 Jun 2023 (this version, v2)]

Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models

Sarah J. Zhang, Samuel Florin, Ariel N. Lee, Eamon Niknafs, Andrei Marginean, Annie Wang, Keith Tyser, Zad Chin, Yann Hicke, Nikhil Singh, Madeleine Udell, Yoon Kim, Tonio Buonassisi, Armando Solar-Lezama, Iddo Drori

We curate a comprehensive dataset of 4,550 questions and solutions from problem sets, midterm exams, and final exams across all MIT Mathematics and Electrical Engineering and Computer Science (EECS) courses required for obtaining a degree. We evaluate the ability of large language models to fulfill the graduation requirements for any MIT major in Mathematics and EECS. Our results demonstrate that GPT-3.5 successfully solves a third of the entire MIT curriculum, while GPT-4, with prompt engineering, achieves a perfect solve rate on a test set excluding questions based on images. We fine-tune an open-source large language model on this dataset. We employ GPT-4 to automatically grade model responses, providing a detailed performance breakdown by course, question, and answer type. By embedding questions in a low-dimensional space, we explore the relationships between questions, topics, and classes and discover which questions and classes are required for solving other questions and classes through few-shot learning. Our analysis offers valuable insights into course prerequisites and curriculum design, highlighting language models' potential for learning and improving Mathematics and EECS education.

Comments: Did not receive permission to release the data or model fine-tuned on the data

Subjects: Computation and Language (cs.CL); Artificial Intelligence (cs.Al); Machine Learning (cs.LG)

arXiv:2306.08997 [cs.CL]

(or arXiv:2306.08997v2 [cs.CL] for this version) https://doi.org/10.48550/arXiv.2306.08997

Submission history

From: Iddo Drori [view email]

[v1] Thu, 15 Jun 2023 09:48:14 UTC (10,482 KB)



Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models

Presents a comprehensive dataset of 4,550 questions and solutions from all MIT EECS courses required for obtaining a degree

arxiv.org/abs/2306.08997

Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models

Sarah J. Zhang*

Sam Florin* MIT sflorin@mit.edu

sjzhang@mit.edu

Ariel N. Lee Boston University ariellee@bu.edu

Eamon Niknafs Andrei Marginean Boston University atmargi@mit.edu annewang@mit.edu

Annie Wang

Keith Tyser Boston University ktyser@bu.edu

Zad Chin Harvard University zadchin@college.harvard.edu

Yann Hicke Cornell University vlh8@cornell.edu

Nikhil Singh MIT nsingh1@mit.edu udell@stanford.edu voonkim@mit.edu buonassi@mit.edu

Madeleine Udell Stanford University

en@bu.edu

Yoon Kim MIT

Tonio Buonassisi MIT

Armando Solar-Lezama asolar@csail.mit.edu

Iddo Drori MIT, Columbia University, Boston University idrori@csail.mit.edu

10:14 PM - Jun 15 2023 - 1.8M Views

139 Retweets 53 Quotes 624 Likes 398 Bookmarks



Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models

Presents a comprehensive dataset of 4,550 questions and solutions from all MIT EECS courses required for obtaining a degree

arxiv.org/abs/2306.08997

Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models

Sarah J. Zhang* sjzhang@mit.edu

Sam Florin* MIT sflorin@mit.edu

Ariel N. Lee Boston University ariellee@bu.edu

Eamon Niknafs Andrei Marginean Boston University atmargi@mit.edu en@bu.edu

Annie Wang annewang@mit.edu

Keith Tyser Boston University ktyser@bu.edu

Zad Chin Harvard University zadchin@college.harvard.edu

Yann Hicke Cornell University ylh8@cornell.edu

Nikhil Singh nsingh1@mit.edu

Madeleine Udell Stanford University udell@stanford.edu yoonkim@mit.edu

Tonio Buonassisi buonassi@mit.edu

Armando Solar-Lezama asolar@csail.mit.edu

Iddo Drori MIT, Columbia University, Boston University idrori@csail.mit.edu

Yoon Kim

10:14 PM · Jun 15, 2023 · 1.8M Views

139 Retweets 53 Ouotes 624 Likes 398 Bookmarks



A recent work from @iddo claimed GPT4 can score 100% on MIT's EECS curriculum with the right prompting.

My friends and I were excited to read the analysis behind such a feat, but after digging deeper, what we found left us surprised and disappointed.

dub.sh/gptsucksatmit



Aran Komatsuzaki 🐶 @arankomatsuzaki - Jun 15

Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models

Presents a comprehensive dataset of 4,550 questions and solutions from all MIT EECS courses required for obtaining a degree

arxiv.org/abs/2306.08997

Exploring the MIT Mathematics and EECS Curriculum Using Large Language Models

Sarah J. Zhang* sjzhang@mit.edu

Sam Florin* sflorin@mit.edu

Ariel N. Lee Boston University ariellee@bu.edu

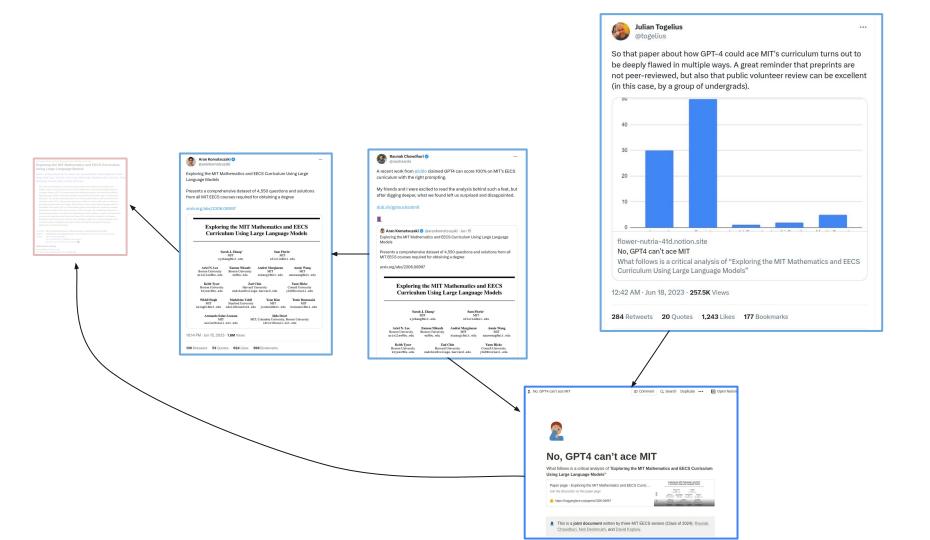
Eamon Niknafs Boston University en@bu.edu

Andrei Marginean Annie Wang atmargi@mit.edu annewang@mit.edu

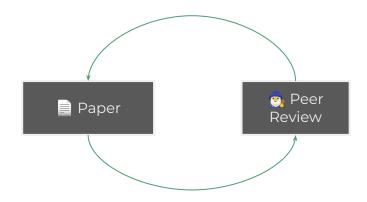
Keith Tyser **Boston University** ktyser@bu.edu

Zad Chin Harvard University zadchin@college.harvard.edu

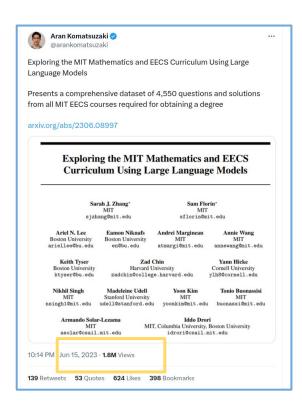
Yann Hicke Cornell University ylh8@cornell.edu ...



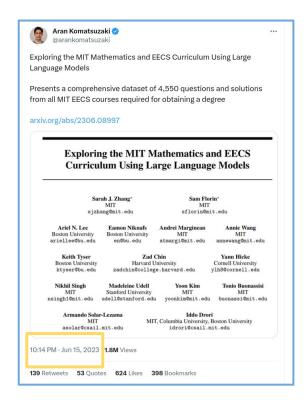
3 problems with traditional science publishing

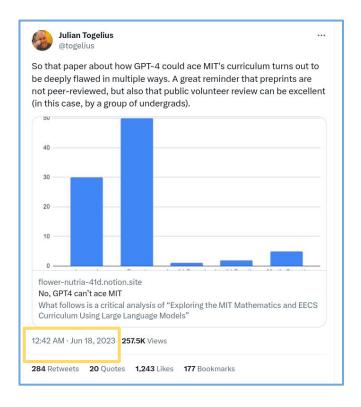


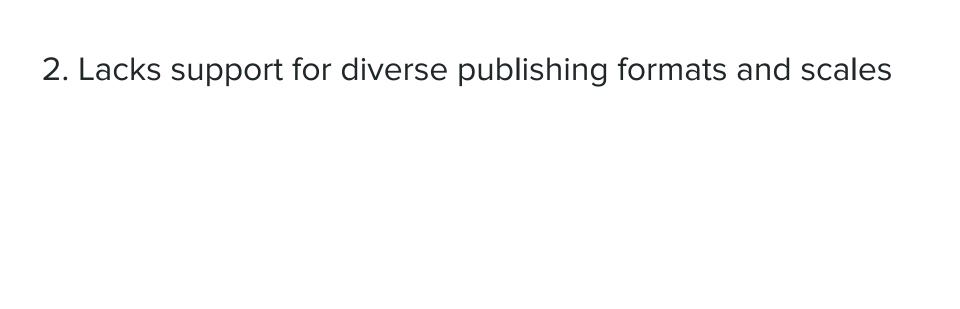
1. Lacks reach + rapid feedback

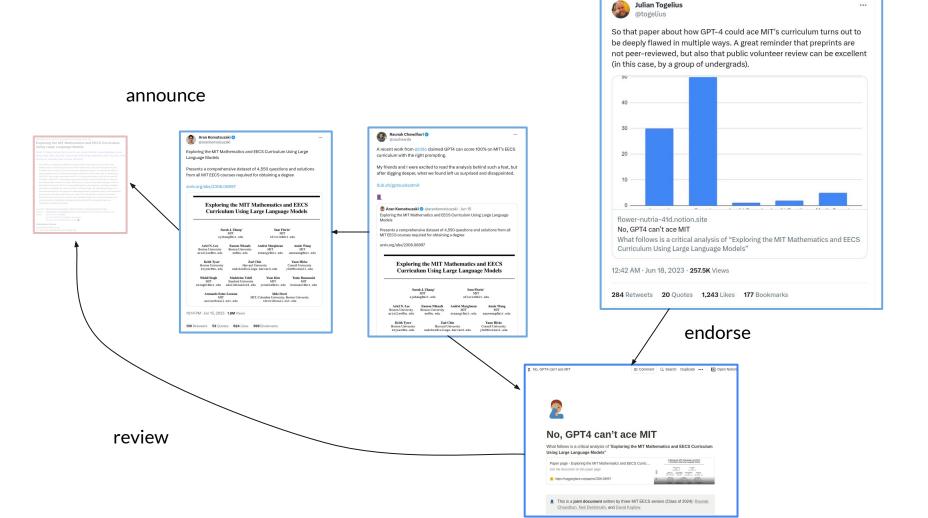


1. Lacks reach + rapid feedback











Looking forward to reading this.

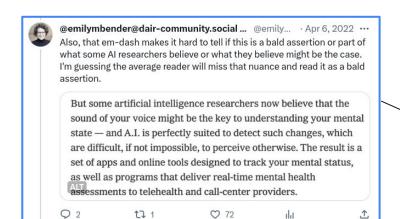


Excited to share the preprint of a letter/opinion paper I wrote with my supervisor Ellen Lau! In this paper we propose a new research program, 'Binding Problem 2.0', extending the question from visual representations to conceptual representations. 1/9 psyarxiv.com/8hj7f

8:51 PM · Nov 2, 2022

reading-status





inline-annotation



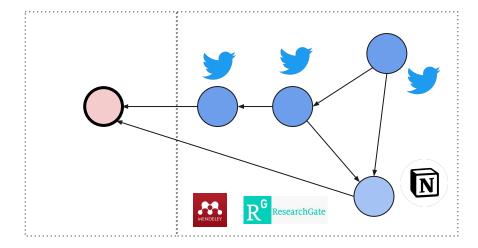
3. Data siloed/fragmented across multiple apps & formats

"altmetrics have been hamstrung by a lack of open access to the event streams we want to analyze...Most APIs are heavily restricted or very expensive to use at scale (e.g., Twitter) or they just don't exist: ever-greater swaths of the scholarly conversation now disappear into the profiteering maw of ResearchGate and other walled gardens, never to be seen again."

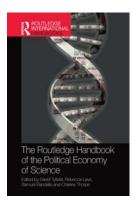
The state of altmetrics (2019)

Paper

Sensemaking about paper

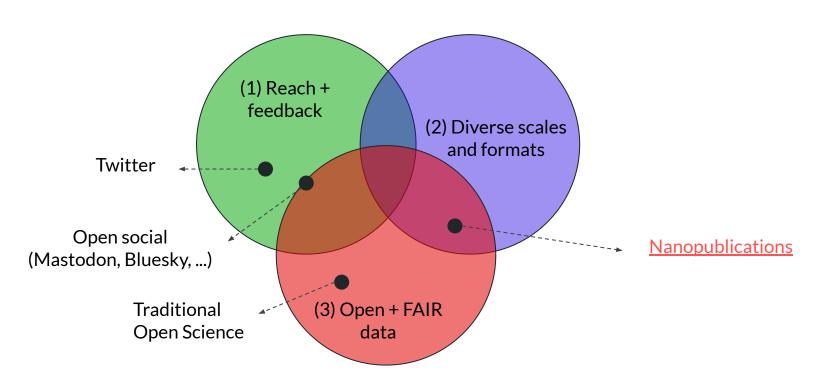


Scientific record, Open Access Off scientific record, controlled by commercial Platforms



Open Access Panacea (Muellerleile, 2017)

"...focusing too closely on ... openness may be distracting us from the ways that capital is sneaking in the back door and enclosing the **very tools we need to make sense** of this new world"



nanodash, beta by knowledge | my channel | users | connectors | search | pu

Publish a new Nanopublication

Choose a nanopublication template below.



FAIR

- Declaring the completion of a 3PFF Programme by Barbara Magagna, 2023-10-15
- Defining a FAIR implementation community by Barbara Magagna, 2023-10-15
- Defining a 3PFF event by Barbara Magagna, 2023-10-07
- Defining a Validation Service by Barbara Magagna, 2023-10-04
- Declaring a FAIR Qualification Criteria by Barbara Magagna, 2023-10-04
- Declaring a FAIR Interpretation by Barbara Magagna, 2023-10-04
- Defining a Registry by Barbara Magagna, 2023-09-26
- . Introducing a new SIP related term (new ID) by Barbara Magagna, 2023-09-16
- Defining a Crosswalk by Barbara Magagna, 2023-09-15
- Defining a Web API by Barbara Magagna, 2023-09-15
- . Defining a FAIR practice by Barbara Magagna, 2023-09-14
- . Defining a Provenance tracking service by Barbara Magagna, 2023-09-14
- Defining a FAIR representation service by Barbara Magagna, 2023-09-14
- Defining a Metadata Schema by Barbara Magagna, 2023-09-12
- Defining a Structured Vocabulary by Barbara Magagna, 2023-09-12
- Defining a Semantic Model by Barbara Magagna, 2023-09-12
- . Defining an Editor by Barbara Magagna, 2023-09-12
- . Introducing a Digital Object Type or Methodology (existing ID) by Tobias Kuhn, 2023-09-01
- Introducing a Digital Object Type or Methodology (new ID) by Tobias Kuhn, 2023-09-01

Social

- . Stating an opinion that something is over-/underrated by Tobias Kuhn, 2023-10-03
- . Defining a group of people by Tobias Kuhn, 2023-08-16
- Announcing a paper that I have read by Tobias Kuhn, 2023-08-16

Publish a new Nanopublication

Assertion: Announcing a paper that I have read ^

I (Ronen Tamari) have read the paper the paper with DOI			
DOI for the paper starting with '10.'			
The paper with DOI DOI for the paper starting with '10.'			
has my comment " comment text	п		

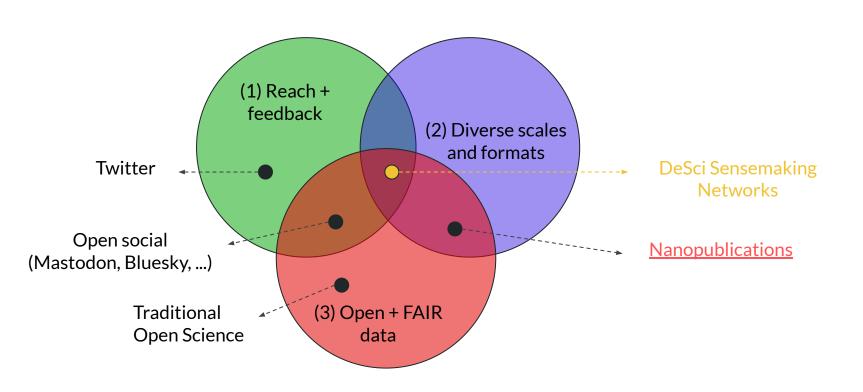
Publish a new Nanopublication

Assertion: Making a review comment according to the LinkFlows model ^

		comment		review comment . 5 to identifier of the object that is reviewed (e.g. a text or nanopublication)	
	This	comment	refer	rs to the mentioning of dentifier of a thing mentioned in the object that is reviewed	ptional) +
۱	This	comment	is a	comment about the content x .	
	This	comment	is a	positive comment x .	
	This	comment	is a	suggestion x v .	
	This	comment	has im	mpact level " 4 " .	
	This	comment	has th	he comment text "The content of this comment as free text " .	

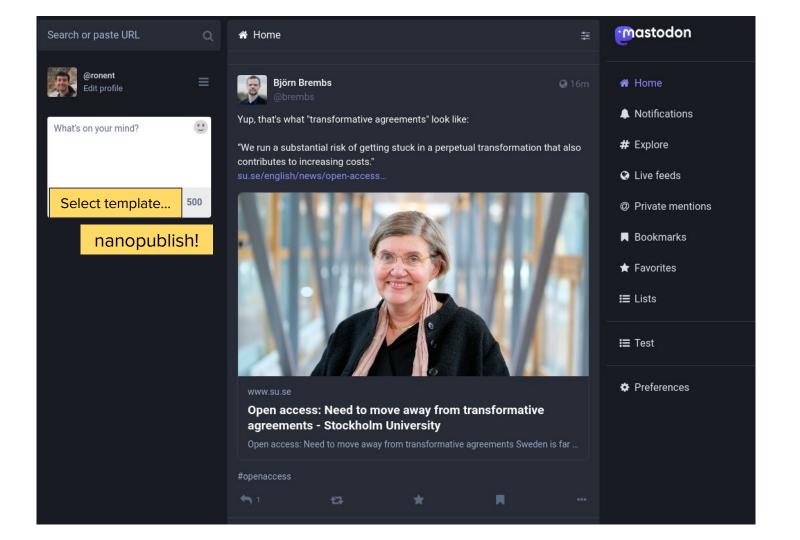
Nanopublications

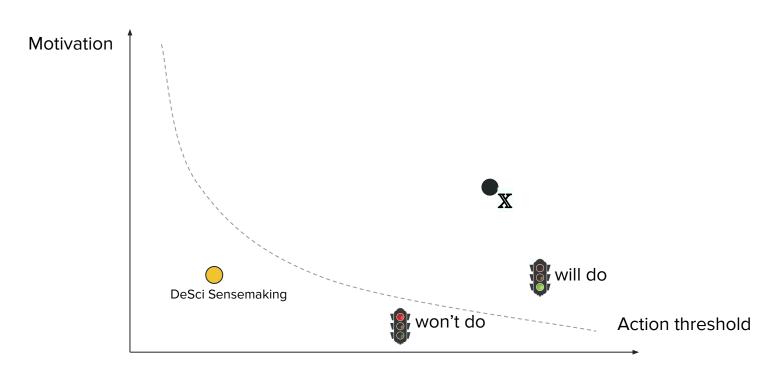
- Human authors as authoritative source for scientific knowledge (not algos!)
- "Text mining? ...Why bury it first and then mine it again?" Barend Mons



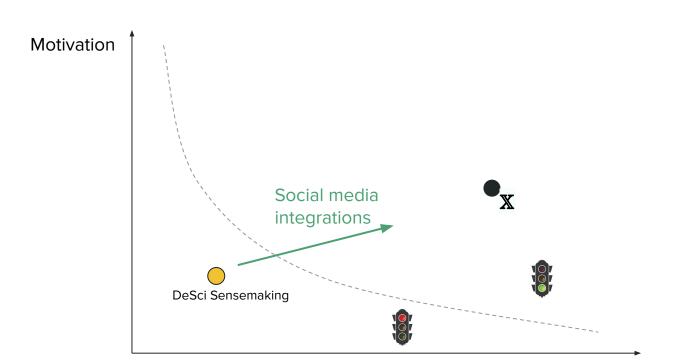
DeSci Sensemaking Networks

- Twitter-style social network
- Posts are "sensemarks": structured semantic objects (nanopubs) with interaction affordances (share, like, etc)
- Open + FAIR data
- Configurable algorithmic feeds
- ID: ORCID or similar web3 variant (OpSci)

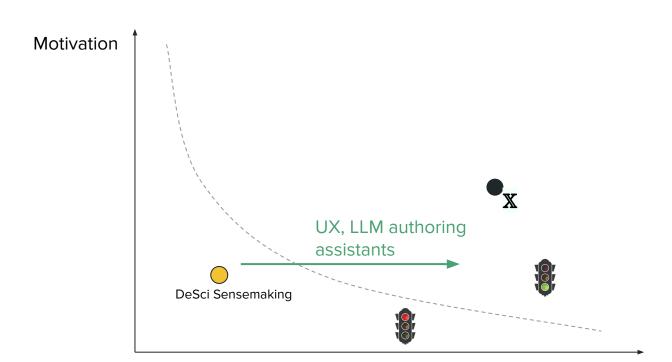




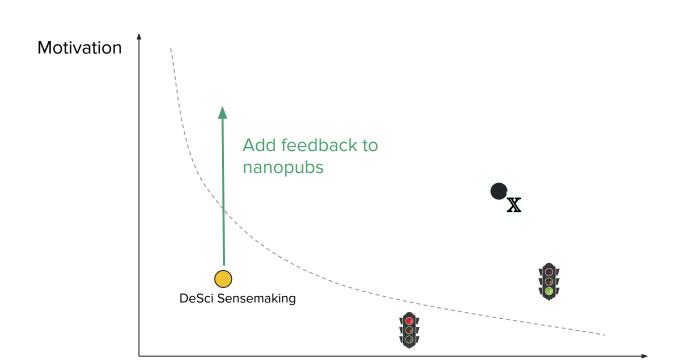
Ease of use



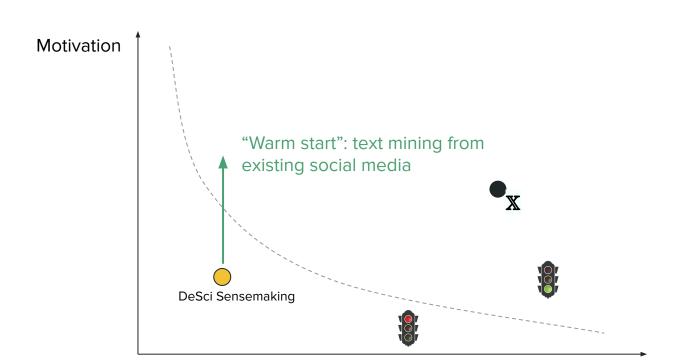
Ease of use



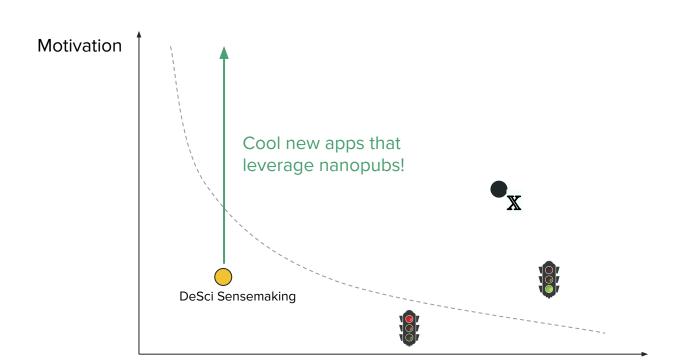
Ease of use



Ease of use



Ease of use



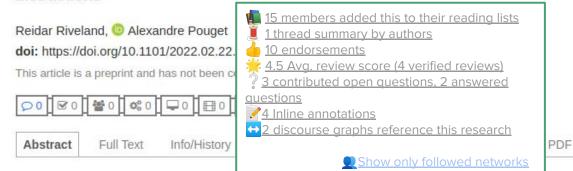
Ease of use

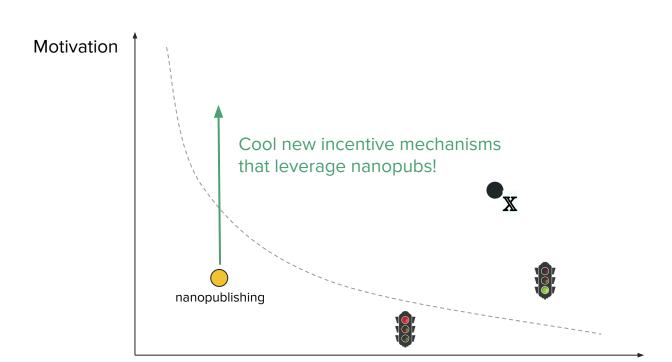




New Results
A Follow this preprint

Generalization in Sensorimotor Networks Configured with Natural Language Instructions

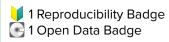




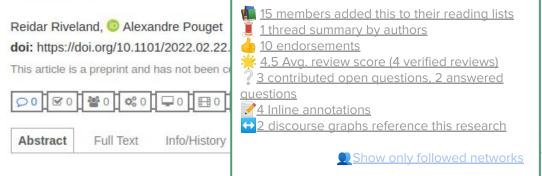
Ease of use



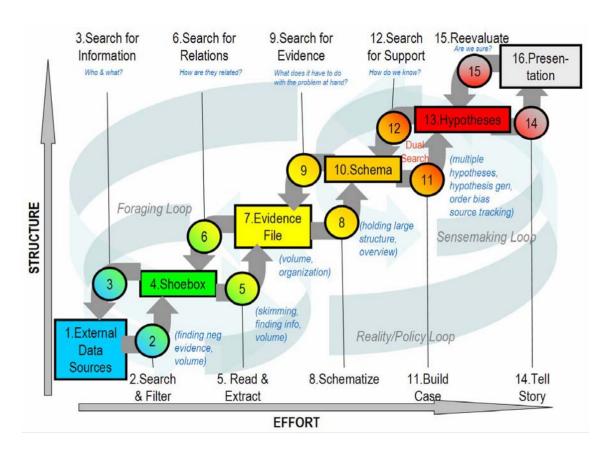




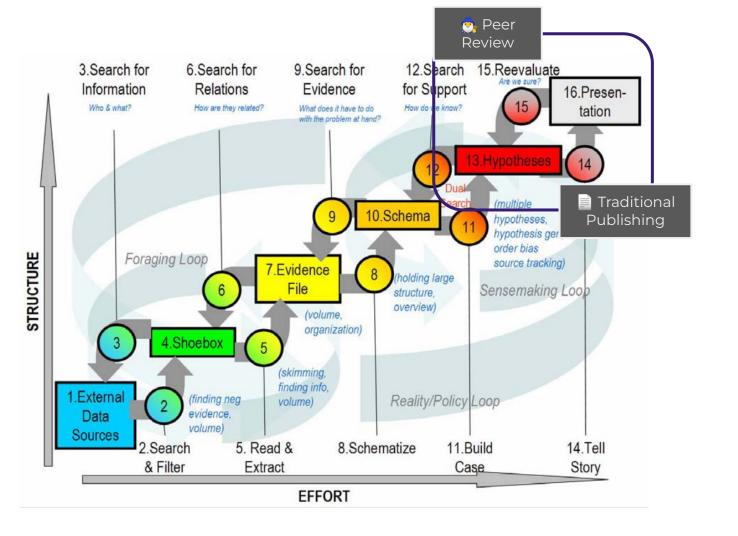
Generalization in Sensorimotor Networks Configured with Natural Language Instructions

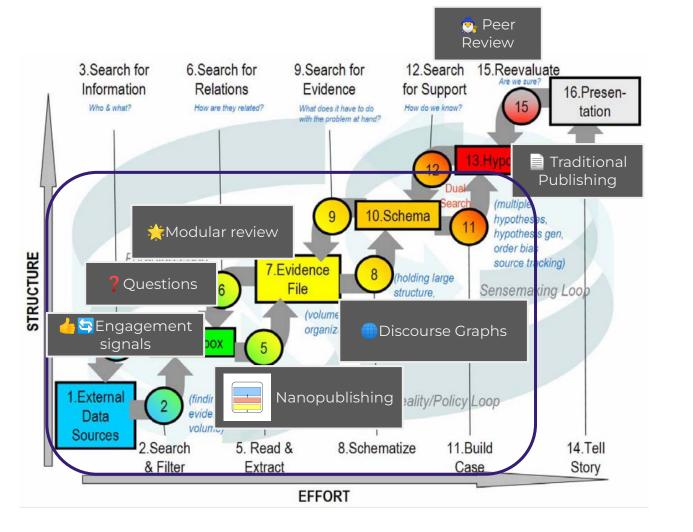


PDF



Source: Pirolli, P., & Card, S. (2005)





DeSci Sensemaking as a sociotechnical movement

Traditional Science Sensemaking	DeSci Sensemaking		
Power Centralization	Power Decentralization		
Exclusive, monoculture participation ecosystem	Inclusive, diverse participation ecosystem		
Neoliberal science - driven by competition	Regen science - driven by cooperation		
Knowledge fragmentation , data enclosure	Knowledge integration, data sharing		

Calling fellow SenseMakers!

Collective sensemaking is... collective!

Join us if you're interested in contributing to {research, development, community building, anything else

