VIII Международный командно-личный турнир школьников по математическому моделированию (ТММ 2025)

The 8th International Team Mathematical Modelling Tournament for High-School Students (MMT-2025)

Rushan Ziatdinov
Professor at Keimyung University, Daegu, South Korea
https://www.ziatdinov-lab.com/
ziatdinov@kmu.ac.kr

Jury Member of the MMT-2025

Задание МаМоНТ-2025 (MMC/Mammoth-2025 Problem)

Чума 2.0 (Plague 2.0)

Альтернативные названия, которые рассматривались: ИИ-мор, Моровая ложь, Информационная чума 1.0
Alternative titles that were considered: AI contagion, Factoid plague, Information Plague 1.0

Чума 2.0 (Plague 2.0)

- Сейчас всё больше и больше текстов, видео, музыки и прочего контента в интернете генерируется при помощи различных систем искусственного интеллекта (ИИ), включая большие языковые модели (LLM, large language models).
- Определенная часть сгенерированных материалов оказывается результатом так называемых **«галлюцинаций» ИИ**, то есть содержит ложную информацию, при этом подаваемую с серьёзным тоном и с претензией на авторитетность источника.
- В сетевом сообществе такие материалы получили название «AI slop», что можно перевести как **«ИИ-помои»**.
- Причины распространения этих «помоев» понятны: если созданные без вложения человеческих усилий материалы удерживают внимание аудитории достаточно долго, то это генерирует «просмотры» сопряжённой с «контентом» рекламы, а значит – и рекламные выплаты авторам контента.

- Nowadays, more and more texts, videos, music, and other content on the internet is generated using various artificial intelligence (AI) systems, including large language models (LLMs).
- A certain portion of the generated material is the result of so-called **AI “hallucinations”**, i.e. it contains false information presented in a serious tone and with a claim to authority.
- The online community has dubbed such materials **“AI slop”**.
- The reasons for the spread of this “slop” are clear: sufficiently long attention of the audience to materials created without any human effort generates “views” of advertisements associated with this content, and the content authors receive advertising payments.

ВикипедиЯ
Свободная энциклопедия

Заглавная страница
Содержание
Избранные статьи
Случайная статья
Текущие события
Пожертвовать

Участие
Сообщить об ошибке
Как править статьи
Сообщество
Форум
Справка
Свежие правки
Новые страницы
Служебные страницы

Инструменты
Ссылки сюда
Связанные правки
Постоянная ссылка
Сведения о странице
Цитировать страницу
Получить короткий URL
Скачать QR-код

Печать/экспорт
Скачать как PDF
Версия для печати

В других проектах
Викисклад
Элемент Викиданных

На других языках
العربية
Deutsch
English

Галлюцинация (искусственный интеллект)

Материал из Википедии — свободной энциклопедии [править код]

У этого термина существуют и другие значения, см. Галлюцинация (значения).

Галлюцинация — уверенная реакция системы искусственного интеллекта, не соответствующая исходным данным, на которых проводилось её обучение; вымышленные ответы системы, не имеющие отношения к действительности[1].

Например, чат-бот, не получив в исходных данных сведений о доходах Tesla, может выбрать случайное число (например, «$13,6 млрд»), которое включит в собственную систему знаний и в дальнейшем будет неоднократно настаивать на этом без каких-либо признаков критического пересмотра[2].

Широкое внимание к подобного рода эффектам проявилось в 2022 году в связи с массовым внедрением больших языковых моделей: пользователи жаловались, что такие боты часто казались «социопатическими» и бессмысленно встраивали правдоподобно звучащие ложные утверждения в ответы[3][4]; в 2023 году аналитики отнесли галлюцинации к одной из серьёзных проблем технологии больших языковых моделей[5]; в связи с этим Илон Маск и Стив Возняк обратились с открытым письмом с предложением приостановить развитие программ искусственного интеллекта, их обращение поддержали более тысячи предпринимателей и экспертов в отрасли[1].

Ряд исследователей считает, что некоторые ответы систем искусственного интеллекта, классифицируемые пользователями как галлюцинации, могут быть объяснены неспособностью пользователей оценить правильность ответа, в качестве примера приводилось точечное изображение, выглядящее для человека как обычное изображение собаки, но на котором программная система обнаруживает фрагменты узора шерсти, характерные только для кошек, ставя высокий приоритет признакам, к которым люди нечувствительны[6]; эти выводы были оспорены другими исследователями[7].

Примечания [править | править код]

1. ↑ 1 2 Искусственный интеллект как угроза человечеству. Маск и Возняк призвали сделать паузу в развитии технологий ИИ. *BBC News Русская служба*. Архивировано 31 марта 2023. Дата обращения: 31 марта 2023.
2. ↑ fastcompany. Дата обращения: 25 марта 2023. Архивировано 29 марта 2023 года.
3. ↑ *Charles Seife*. The Alarming Deceptions at the Heart of an Astounding New Chatbot (амер. англ.) // Slate. — 2022-12-13. — ISSN 1091-2339. Архивировано 26 марта 2023 года.
4. ↑ *Lance Eliot*. AI Ethics Lucidly Questioning This Whole Hallucinating AI Popularized Trend That Has Got To Stop (англ.). *Forbes*. Дата обращения: 25 марта 2023. Архивировано 11 апреля 2023 года.
5. ↑ *Kif Leswing*. Microsoft's Bing A.I. made several factual errors in last week's launch demo (англ.). *CNBC*. Дата обращения: 25 марта 2023. Архивировано 16 февраля 2023 года.
6. ↑ *Louise Matsakis*. Artificial Intelligence May Not 'Hallucinate' After All (амер. англ.) // Wired. — ISSN 1059-1028. Архивировано 26 марта 2023 года.
7. ↑ *Justin Gilmer, Dan Hendrycks*. A Discussion of 'Adversarial Examples Are Not Bugs, They Are Features': Adversarial Example Researchers Need to Expand What is Meant by 'Robustness' (англ.) // Distill. — 2019-08-06. — Vol. 4, iss. 8. — P. e00019.1. — ISSN 2476-0757. — doi:10.23915/distill.00019.1. Архивировано 26 марта 2023 года.

Категории: Языковое моделирование | Дезинформация | Генеративный искусственный интеллект

Hallucination (artificial intelligence)

28 languages

From Wikipedia, the free encyclopedia

This article is about the phenomenon of AI presenting fabricated information as fact. For the appearance of AI-induced psychosis in humans, see Chatbot psychosis.
Not to be confused with Artificial imagination.

In the field of artificial intelligence (AI), a **hallucination** or **artificial hallucination** (also called **bullshitting**,[1][2][3] **confabulation**,[4] or **delusion**[5]) is a response generated by AI that contains false or misleading information presented as fact.[6][7] This term draws a loose analogy with human psychology, where a hallucination typically involves false *percepts*. However, there is a key difference: AI hallucination is associated with erroneously constructed responses (confabulation), rather than perceptual experiences.[7]

For example, a chatbot powered by large language models (LLMs), like ChatGPT, may embed plausible-sounding random falsehoods within its generated content. Detecting and mitigating errors and hallucinations pose significant challenges for practical deployment and reliability of LLMs in high-stakes scenarios, such as chip design, supply chain logistics, and medical diagnostics.[8][9][10] Software engineers and statisticians have criticized the specific term "AI hallucination" for unreasonably anthropomorphizing computers.[11][12]

First-generation Sora video of the Glenfinnan Viaduct in Scotland, incorrectly showing: a second track, trains traveling on the right instead of the left, a second chimney on its interpretation of the train *The Jacobite*, and some carriages much longer than others


The real Glenfinnan Viaduct with *The Jacobite* on it

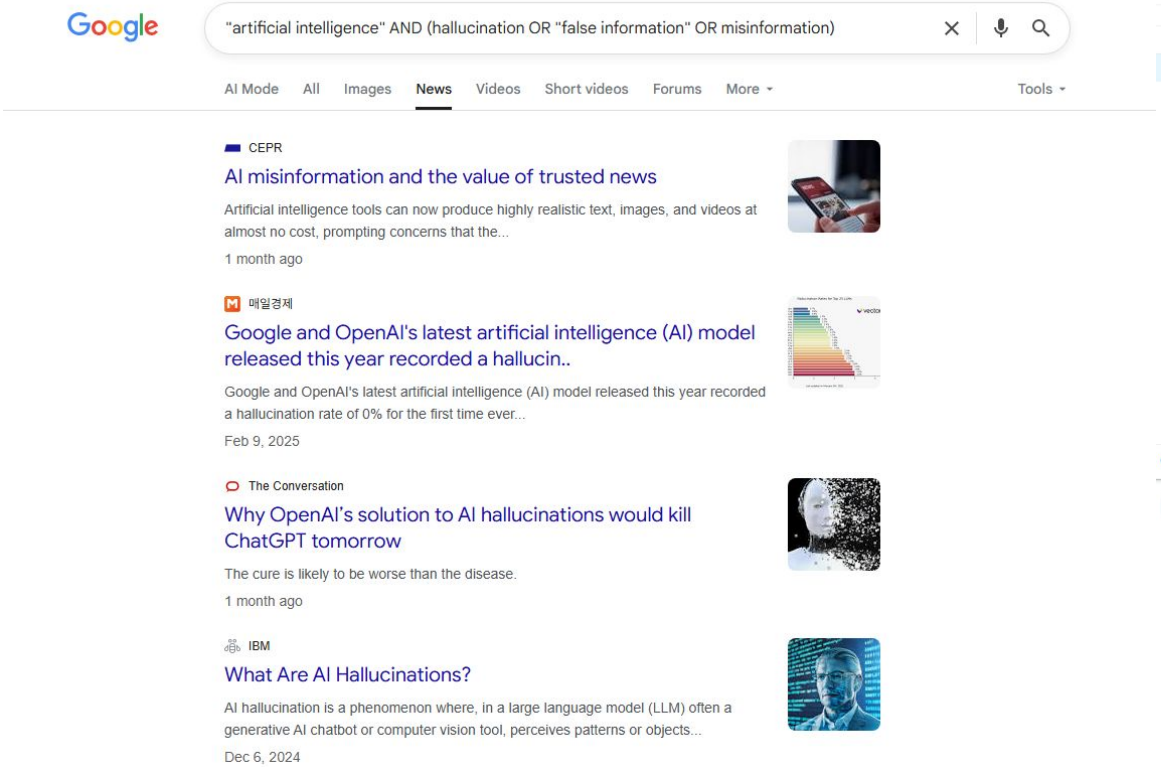
Term [edit]

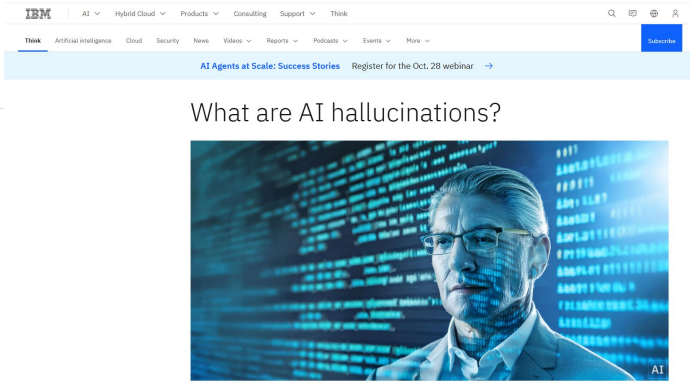
Origin [edit]

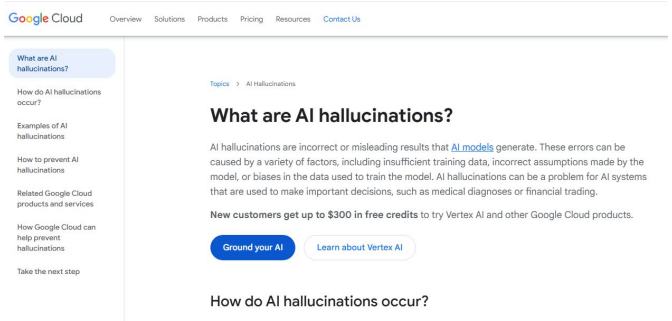
In 1995, Stephen Thaler demonstrated how hallucinations and phantom experiences emerge from artificial neural networks through random perturbation of their connection weights.[a]

In the early 2000s, the term "hallucination" was used in computer vision with a positive connotation to describe the process of adding detail to an image. For example, the task of generating high-resolution face images from low-resolution inputs is called face hallucination.[18][19]

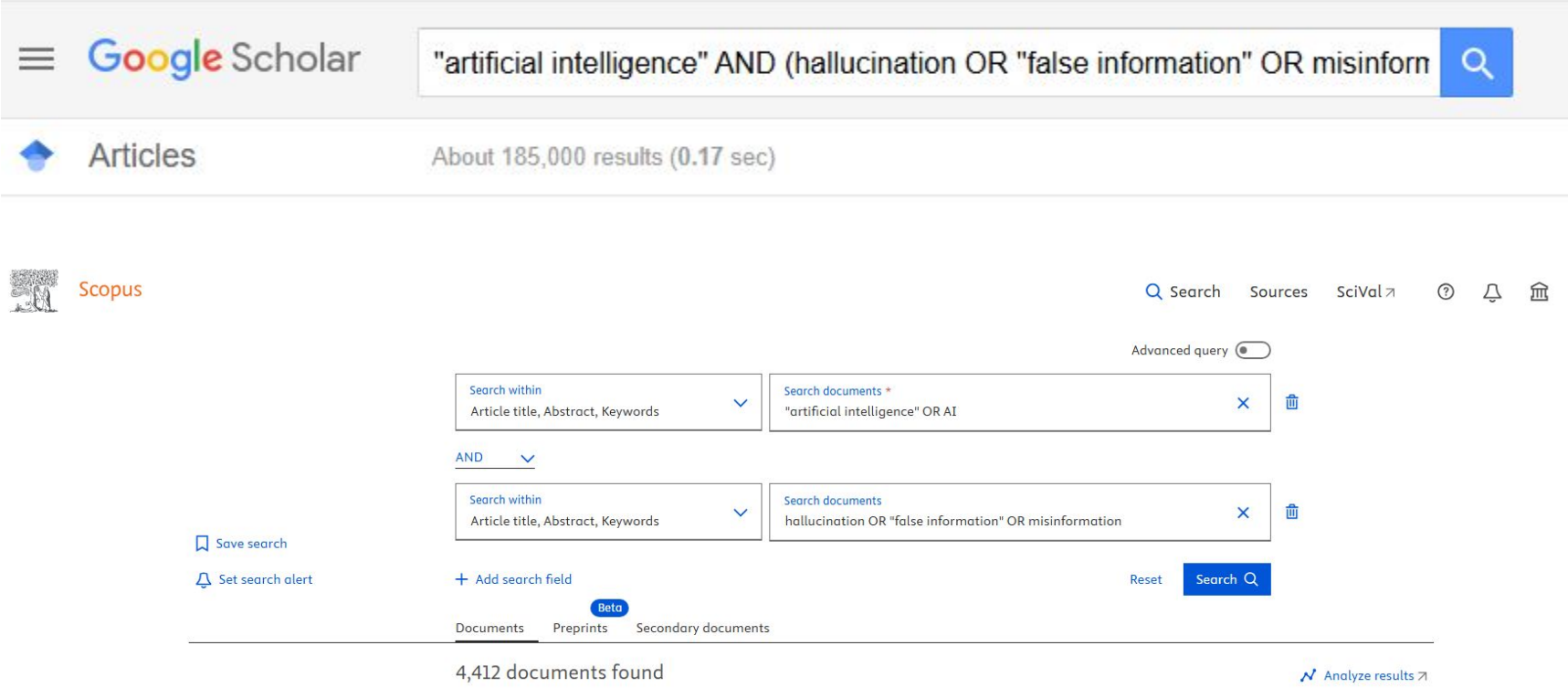
News


Google
"artificial intelligence" AND (hallucination OR "false information" OR misinformation)
AI Mode
All
Images
News
Videos
Short videos
Forums
More
Tools
CEPR
AI misinformation and the value of trusted news
Artificial intelligence tools can now produce highly realistic text, images, and videos at almost no cost, prompting concerns that the...
1 month ago
매일경제
Google and OpenAI's latest artificial intelligence (AI) model released this year recorded a hallucin..
Google and OpenAI's latest artificial intelligence (AI) model released this year recorded a hallucination rate of 0% for the first time ever...
Feb 9, 2025
The Conversation
Why OpenAI's solution to AI hallucinations would kill ChatGPT tomorrow
The cure is likely to be worse than the disease.
1 month ago
IBM
What Are AI Hallucinations?
AI hallucination is a phenomenon where, in a large language model (LLM) often a generative AI chatbot or computer vision tool, perceives patterns or objects...
Dec 6, 2024


IBM
AI AI Agents at Scale: Success Stories
Register for the Oct. 28 webinar
What are AI hallucinations?


Google Cloud
Overview
Solutions
Products
Pricing
Resources
Contact Us
What are AI hallucinations?
How do AI hallucinations occur?
Examples of AI hallucinations
How to prevent AI hallucinations
Related Google Cloud products and services
How Google Cloud can help prevent hallucinations
Take the next step
Topics > AI Hallucinations
What are AI hallucinations?
AI hallucinations are incorrect or misleading results that AI models generate. These errors can be caused by a variety of factors, including insufficient training data, incorrect assumptions made by the model, or biases in the data used to train the model. AI hallucinations can be a problem for AI systems that are used to make important decisions, such as medical diagnoses or financial trading.
New customers get up to $300 in free credits to try Vertex AI and other Google Cloud products.
Ground your AI
Learn about Vertex AI
How do AI hallucinations occur?

Google Scholar & Scopus Search



Чума 2.0 (Plague 2.0)

- Однако у этого явления существует и обратная сторона: развитие ИИ продолжается, исследователи продолжают создавать и тренировать новые ИИ-модели, которые становятся всё более и более крупными, а потому требуют всё большего и большего объёма информации для обучения.
- Практически единственным достаточно крупным источником данных является содержимое интернета.
- **А значит, новые ИИ-модели могут находить «значимые» материалы (что измеряется бо́льши́м количеством «просмотров»), на самом деле являющиеся теми самыми «ИИ-помоями», и использовать их для своего обучения, считая их истиной по определению.**
- В результате этого новые ИИ-модели могут всё более и более отдаляться от фактов реального мира.

- However, this phenomenon has the flip side: AI developers do not stop, they create new, larger AI models, which need more and more information for training.
- But in practice, there is only one source of data large enough to satisfy the growing needs of AI training: the internet content.
- **This means that new AI models can find “significant” materials (where “significance” is measured by the larger number of “views”), which are, in fact, “AI slop” themselves, consider them to be true by definition, and use them for their training.**
- As a result, new AI models may stray further and further from the facts of the real world.

Каковы перспективы этого процесса? Через какое время интернет «потонет» в «ИИ-помоях»? Что может повлиять на этот процесс?

What are the prospects of this process? How long will it take for the internet to be “flooded” with “AI slop”? What factors could influence this process?

Задачи (Tasks)

1. Используя **открытые источники и статистические данные**, оцените:

1.1. Скорость заражения новых ИИ-моделей ложными фактами («factoids») из интернета: сколько времени может пройти от появления факта в интернет-материалах и до его включения в широко используемые ИИ-модели?

1.2. Какая доля «популярных» интернет-публикаций является сгенерированной ИИ-системами, и какая доля от них скорее всего является ложной информацией, являющейся результатом «галлюцинаций» ИИ? Поясните, какие материалы вы относите к «популярным».

1. Using **open sources and statistical data**, estimate:

1.1. The rate at which new AI models are contaminated by false facts ("factoids") from the internet: how long might it take from a factoid appearing in online materials to its inclusion into widely-used AI models?

1.2. What proportion of “popular” online publications is generated by AI systems, and what proportion of those is likely to be false information resulting from AI “hallucinations”? Please explain which materials you consider to be “popular”.

Задачи (Tasks)

2. Предполагая, что доля ложных фактов в общем количестве интернет-информации изначально мала и остаётся сравнительно малой на протяжении какого-то времени, предложите формулу, приблизительно описывающую долю ложных фактов как функцию времени (на протяжении этого «какого-то времени»). Какое время потребуется для достижения доли в 10% и 20%, если процесс начинается с уровня в 0,1%?

2. Assuming that the proportion of false facts in the total amount of internet information is initially small and remains relatively small for a certain period of time, propose a formula that approximately describes the proportion of false facts as a function of time (within this “certain period”). How long will it take to reach a proportion of 10% and 20% if the process starts at 0.1%?

Задачи (Tasks)

3. Используя наработки из заданий 1 и 2, постройте математическую модель, предсказывающую динамику доли ложных фактов в интернет-информации вне зависимости от того, мала ли эта доля или нет. Используя построенную модель, спрогнозируйте развитие ситуации: если поведение пользователей интернета и создателей новых ИИ-систем не изменится, то, в конечном итоге, заполнится ли интернет практически на 100% ложной информацией или доля ложной информации стабилизируется на каком-то более низком значении? Через какое время это может произойти?

3. Using the findings from tasks 1 and 2, build a mathematical model that predicts the dynamics of the proportion of false facts in online information regardless of the proportion being small or not. Using the model you have built, predict how the situation will develop: if the behaviour of internet users and creators of new AI systems does not change, will the internet eventually become filled with false information almost to 100%, or will the proportion of false information stabilise at some lower value? How long might this take?

Задачи (Tasks)

4. Ожидаете ли вы какое-то изменение поведения пользователей интернета и авторов ИИ в ответ на очевидно высокий уровень «ИИ-помоев» в интернете? Как изменятся предсказания вашей модели, если эти изменения поведения будут включены в неё? Спрогнозируйте реальное поведение доли ложной информации в интернете в ближайшие несколько десятилетий.

4. Do you expect any change in the behaviour of internet users and creators of AI systems in response to the obviously high level of “AI slop” on the internet? How will your model's predictions change if these behavioural changes are incorporated into it? Predict the actual behaviour of the proportion of false information on the internet over the next few decades.

Задачи (Tasks)

Заметим, что, очевидно, существуют и другие источники информации для обучения ИИ-систем помимо интернет-материалов – например, научные публикации (статьи) или художественная литература. **Однако и эти сферы сейчас находятся под серьёзным влиянием материалов, сгенерированных ИИ.** Авторы материалов в этих сферах руководствуются во многом иными соображениями, чем авторы контента, предназначенного специально для публикации в интернете, но итоговый результат качественно похож. К тому же, объем информации в научной и художественной литературе суммарно намного меньше, чем «весь интернет».

It should be noted that, obviously, there are other sources of information for training AI systems besides online materials – for example, scientific publications (articles) or fiction literature. **However, these areas are now also heavily influenced by AI-generated materials.** While the authors of materials in these fields are guided by considerations that are largely different from those of the authors of content intended specifically for publication on the internet, the end result is qualitatively similar. In addition, the total amount of information in scientific and fiction literature is much smaller than that of the “entire internet”.

Dangers of AI in Academic Publishing

NEWS | 12 December 2023

More than 10,000 research papers were retracted in 2023 – a new record

The number of articles being retracted rose sharply this year. Integrity experts say that this is only the tip of the iceberg.

By Richard Van Noorden

The number of retractions issued for research articles in 2023 has passed 10,000 – smashing annual records – as publishers struggle to clean up a slew of sham papers and peer-review fraud. Among large research-producing nations, Saudi Arabia, Pakistan, Russia and China have the highest retraction rates over the past two decades, a *Nature* analysis has found.

Related Articles

AI intensifies fight against 'paper mills' that churn out fake research

Paper-mill detector put to the test in push to stamp out fake science

'Tortured phrases' give away fabricated research papers

Scopus

Search Sources SciVal

Advanced query

Search within: Article title, Abstract, Keywords

Search documents *: "mathematical modeling" OR "mathematical modelling"

Save search

Set search alert

+ Add search field

Reset Search

Documents Preprints (Beta) Secondary documents

111,782 documents found

Analyze results

Refine search

Search within results

Filters

Year

Range Individual

from – to

Author name

- undefined 1,402
- Rajendran, L. 130
- Hayat, T. 104
- Perelson, A.S. 83

All Export Download Citation overview ••• More Show all abstracts Sort by Date (oldest)

	Document title	Authors	Source	Year	Citations
1	Conference Paper **MATHEMATICAL MODELLING OF FLOWS IN NATURAL STREAMS.** Show abstract Locate full-text Related documents	Krishnappan, Bommana G.	Electrochimica Acta, 2, pp. 763–787	1960	0
2	Article • *Open access* **Integer Programming Formulation of Traveling Salesman Problems** Show abstract Locate full-text View at Publisher	Miller, C.E., Zemlin, R.A., Tucker, A.W.	Journal of the ACM Jacm, 7(4), pp. 326–329	1960	1,567
3	Conference Paper **Advanced maintainability techniques for aircraft systems** Show abstract Locate full-text View at Publisher	Voegtlen, H.D., Retterer, B.L.	SAE Technical Papers	1963	0

Discover early research ideas

View preprints published by authors to have an early idea of upcoming research documents.

View 4565 preprints

Filter by country/territory

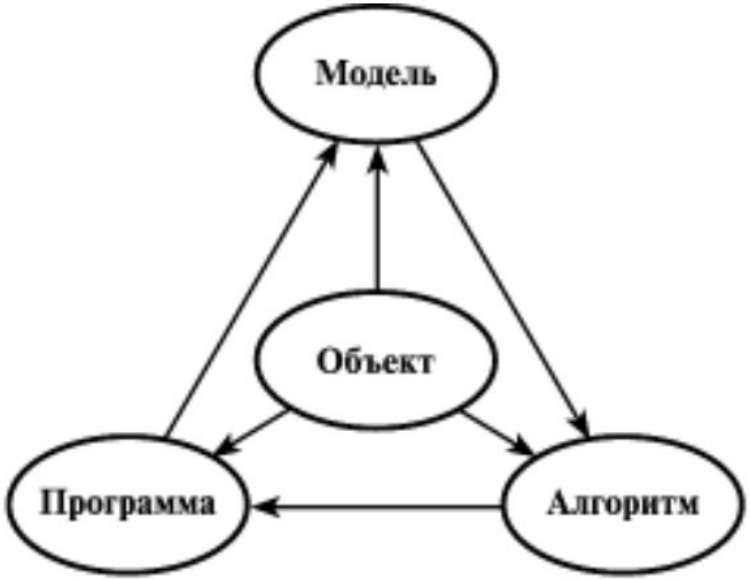
Sort by Number of results

- [ ] United States
- [ ] Russian Federation
- [ ] China
- [ ] India
- [ ] United Kingdom
- [ ] Undefined
- [ ] Germany
- [ ] Ukraine
- [ ] Canada
- [ ] France
- [ ] Italy

Filter by affiliation

Sort by Number of results

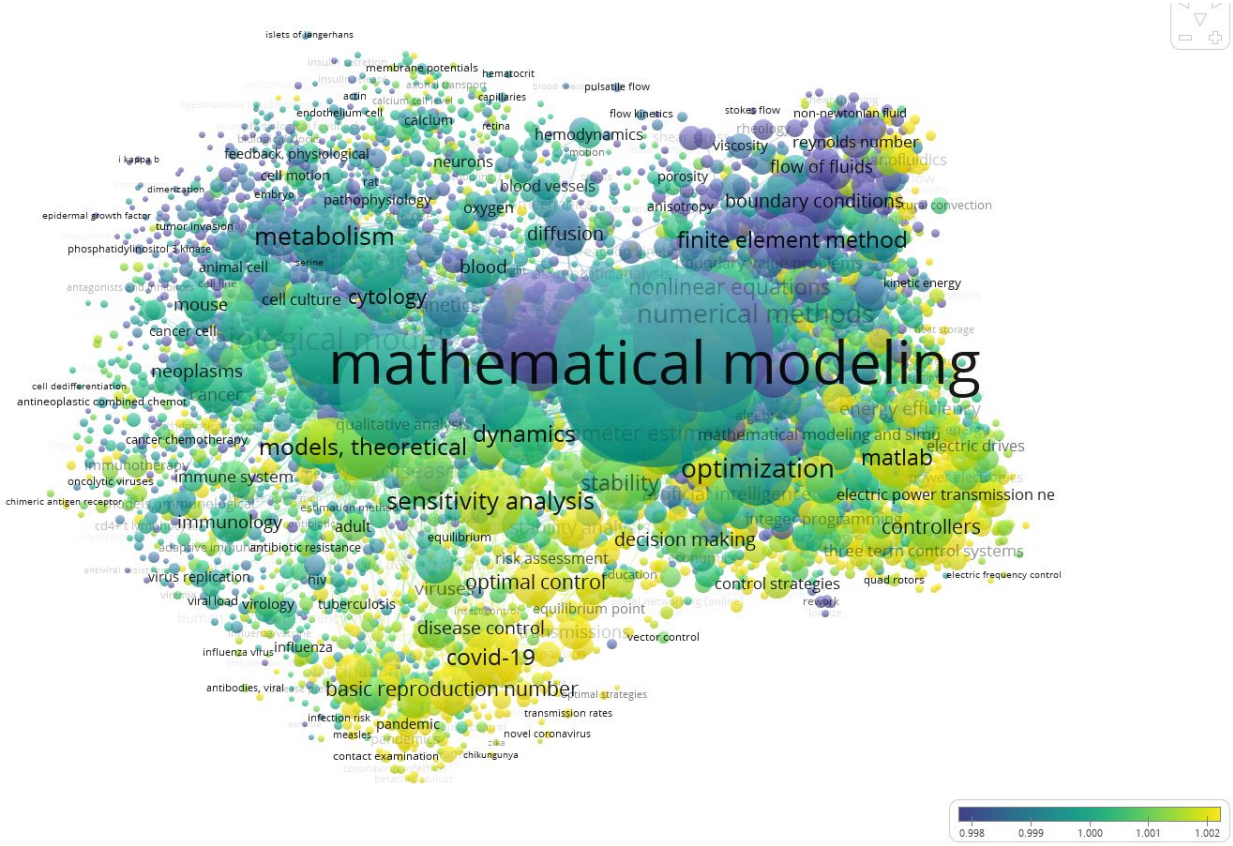
- [ ] Russian Academy of Sciences
- [ ] Siberian Branch, Russian Academy of Sciences
- [ ] Lomonosov Moscow State University
- [ ] Imperial College London
- [ ] CNRS Centre National de la Recherche Scientifique
- [ ] National Academy of Sciences of Ukraine
- [ ] University of Oxford
- [ ] Tomsk Polytechnic University
- [ ] Ministry of Education of the People's Republic of China
- [ ] Chinese Academy of Sciences



«модель – алгоритм – программа» — триада Самарского, отражающая основные этапы математического моделирования (подробнее).

https://samarskii.ru/books/samarskii_book.pdf

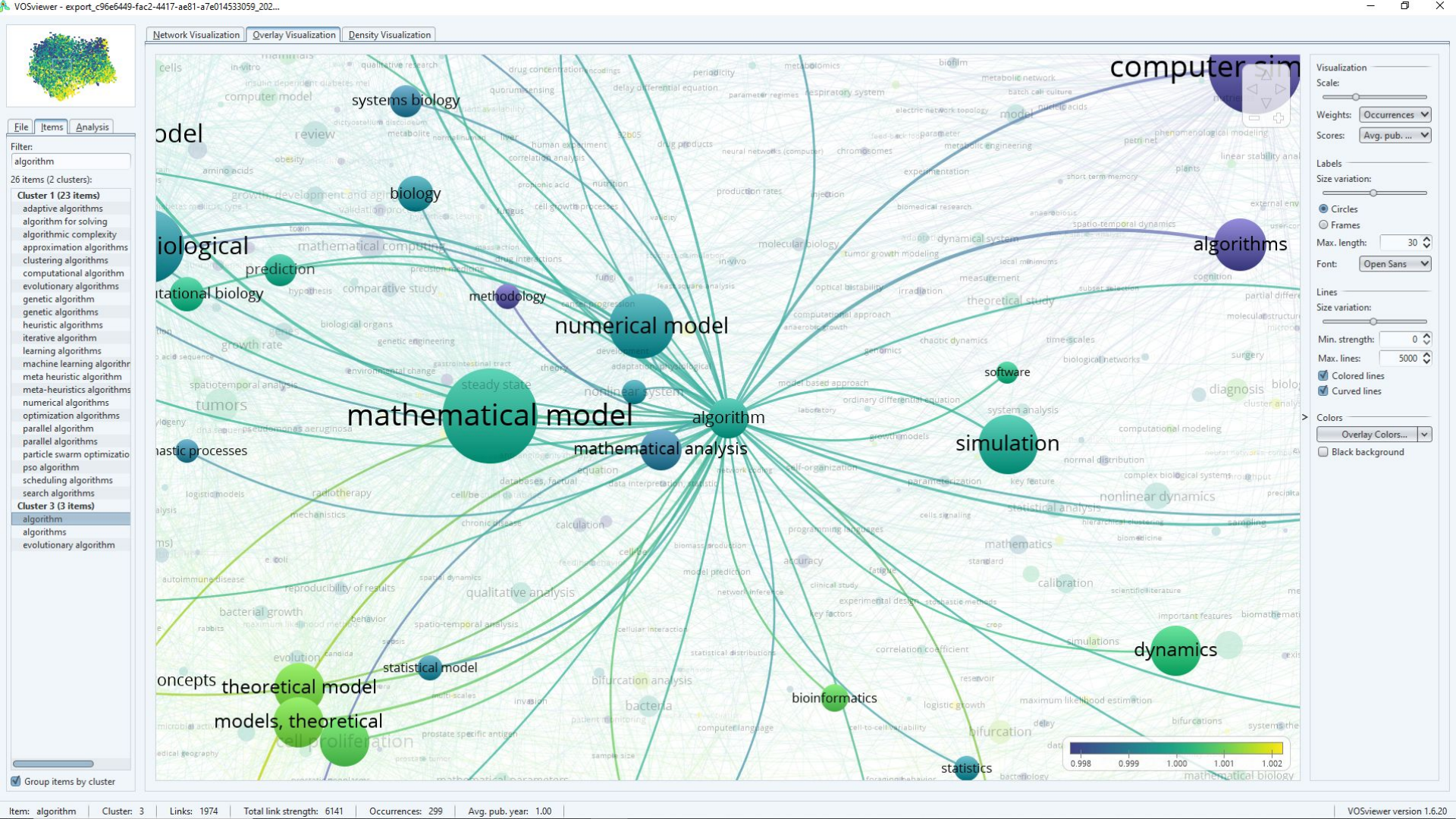
Keyword Network Visualization (20.000 Scopus documents)



Filter by subject area

Sort by Number of results

- [ ] Engineering
- [ ] Mathematics
- [ ] Computer Science
- [ ] Physics and Astronomy
- [ ] Chemical Engineering
- [ ] Materials Science
- [ ] Environmental Science
- [ ] Biochemistry, Genetics and Molecular Biology
- [ ] Chemistry
- [ ] Medicine


VOSviewer - export_c96e6449-fac2-4417-ae81-a7e014533059_202...
Network Visualization
Overlay Visualization
Density Visualization
File
Items
Analysis
Filter:
algorithm
26 items (2 clusters):
Cluster 1 (23 items)
adaptive algorithms
algorithm for solving
algorithmic complexity
approximation algorithms
clustering algorithms
computational algorithm
evolutionary algorithms
genetic algorithm
genetic algorithms
heuristic algorithms
iterative algorithm
learning algorithms
machine learning algorithr
meta heuristic algorithm
meta-heuristics algorithms
numerical algorithms
optimization algorithms
parallel algorithm
parallel algorithms
particle swarm optimizatio
pso algorithm
scheduling algorithms
search algorithms
Cluster 3 (3 items)
algorithm
algorithms
evolutionary algorithm
Group items by cluster
mathematical model
algorithm
mathematical analysis
numerical model
simulation
algorithms
computer simulation
systems biology
biology
methodology
prediction
software
dynamics
bioinformatics
statistical model
theoretical model
models, theoretical
statistics
Visualization
Scale:
Weights:
Occurrences
Scores:
Avg. pub. ...
Labels
Size variation:
Circles
Frames
Max. length:
30
Font:
Open Sans
Lines
Size variation:
Min. strength:
0
Max. lines:
5000
Colored lines
Curved lines
Colors
Overlay Colors...
Black background
0.998
0.999
1.000
1.001
1.002
Item: algorithm
Cluster: 3
Links: 1974
Total link strength: 6141
Occurrences: 299
Avg. pub. year: 1.00
VOSviewer version 1.6.20

File Edit Selection View Go Run Terminal Help

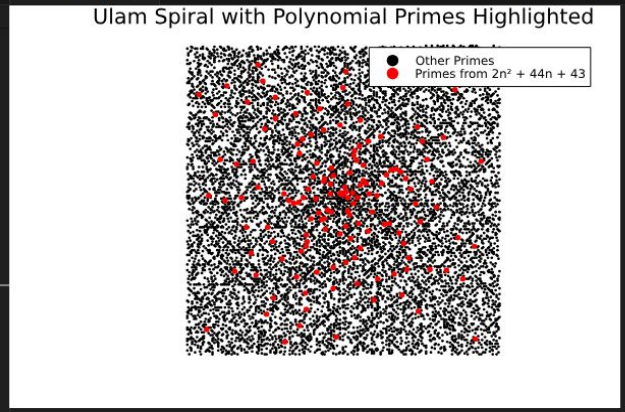
Julia

Ulam Spiral.jl

Ulam Spiral.jl > ...

```
10  function ulam_spiral(n_max)
37              current_step_count = 0
38              steps_at_current_length += 1
39
40              # Increase step length after two turns
41              if steps_at_current_length == 2
42                  steps += 1
43                  steps_at_current_length = 0
44              end
45          end
46      end
47      return coords
48  end
49
50  ## --- Main Script ---
51
52  # 1. Define the maximum number to spiral to (corresponding to the 10,000th prime)
53  # The 10,000th prime is 104729. We need a number slightly larger for the spiral to enc
54  # Let's use 105000 as a safe upper limit for the spiral's range.
55  const N_MAX = 105000
56  const TARGET_PRIME_COUNT = 10000
57
58  # 2. Generate the first 10,000 prime numbers
59  # `primes(n)` is generally more efficient than checking primality for all numbers up t
60  all_primes = primes(N_MAX)
61  first_10k_primes = all_primes[1:min(end, TARGET_PRIME_COUNT)]
62
63  # 3. Define the polynomial and find the numbers it generates
64  poly(n) = 2*n^2 + 44*n + 43
65
66  # Find the maximum 'n' such that poly(n) is still within the N_MAX range and is prime.
67  # We'll check up to a reasonable limit, say n=1 to n=200, or until the polynomial exce
68  max_n_check = 200
69  polynomial_primes = Set{Int}()
70  for n in 1:max_n_check
71      p = poly(n)
72      if p > N_MAX
73          break # Stop if the number exceeds the spiral's range
74      end
75      # Check if p is both prime AND is one of the first 10,000 primes (implicitly covere
76      if isprime(p)
77          push!(polynomial_primes, p)
78      end
```

Julia Plots (1/1)



https://julialang.org/

CHAT

Ask about your code

AI responses may be inaccurate.

Generate Agent Instructions to onboard AI onto your codebase.

PROBLEMS 29 OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
julia>
```

powershell
Run Ula...
Julia REPL (...

Add Context...

Add context (#), extensions (@), commands (/)

Ask Auto

8 0 21 Julia env: v1.12

Main