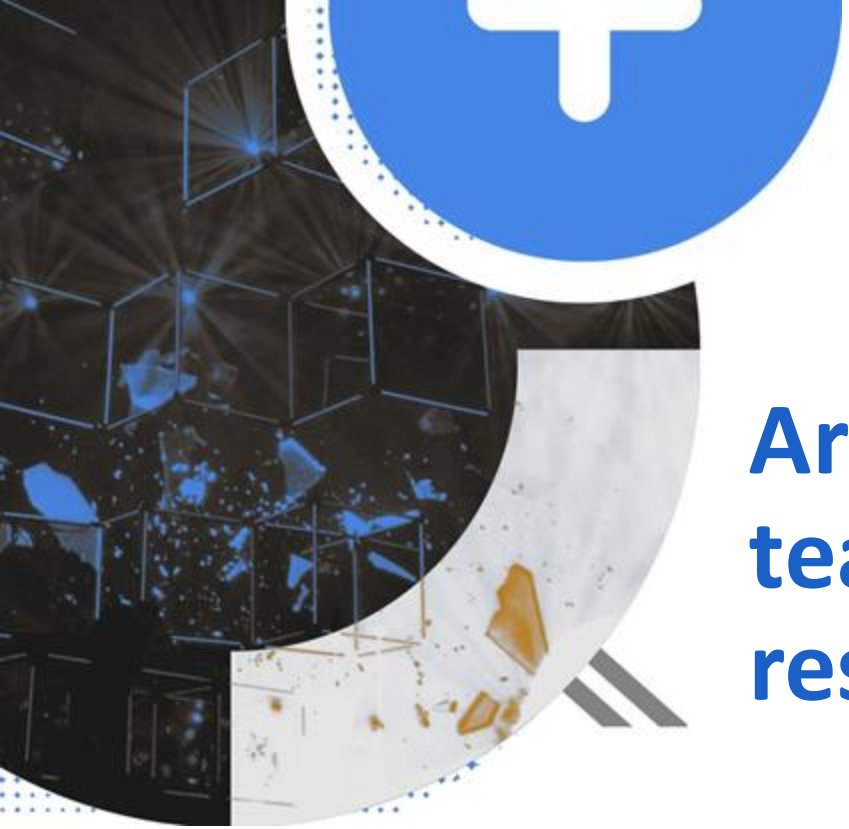


# Artificial intelligence, teaching and research

Jonathan England, OpenAIRE
Milica Ševkušić, EIFL

# Outline

- Definitions
- AI-based tools
- Challenges and recommendations

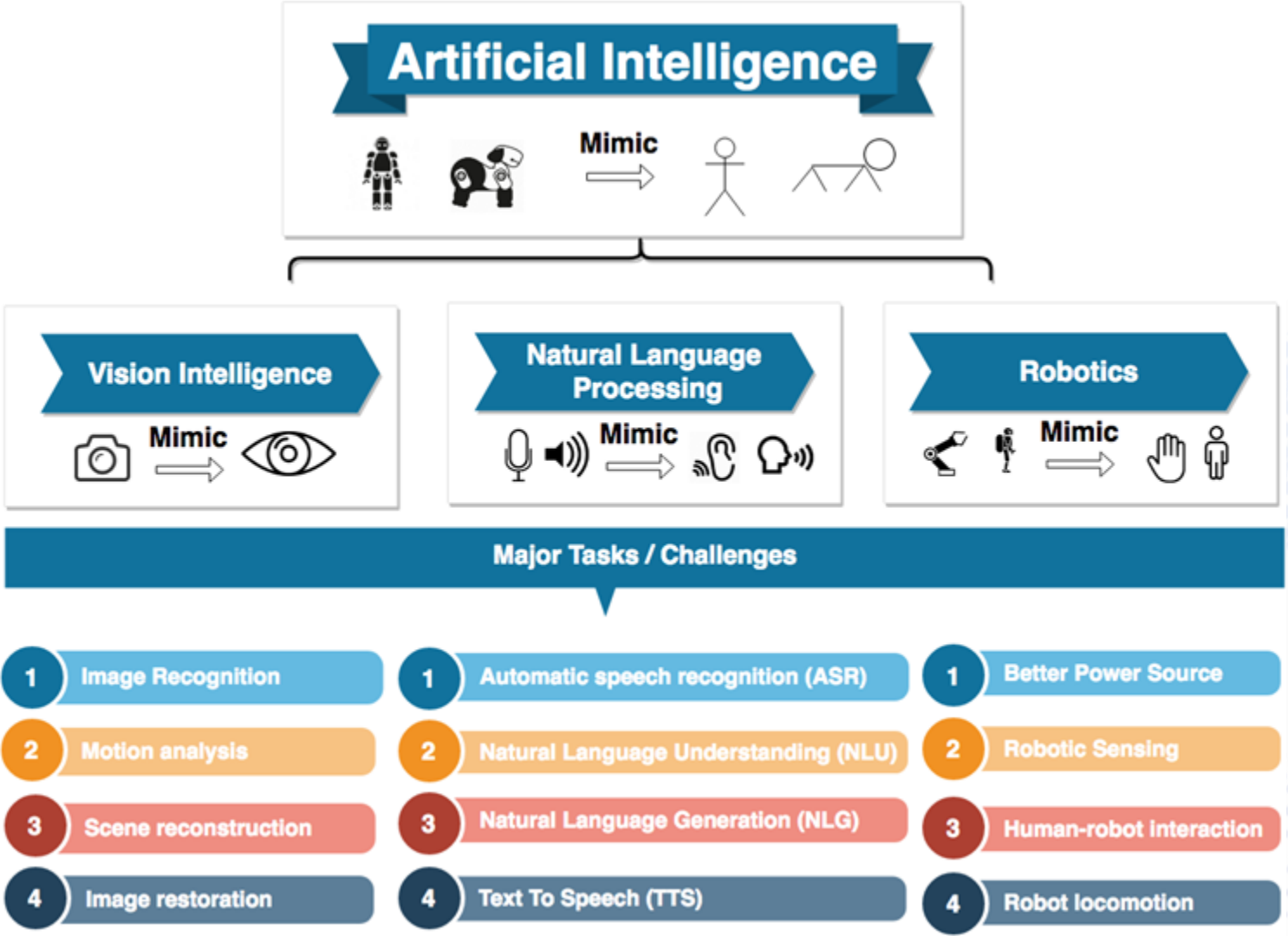
Image by vectorjuice on Freepik

## What is artificial intelligence (AI)?

*As technology develops, so too do the ways we define it. There is no single or fixed definition of AI, but there is common agreement that machines based on AI "are potentially capable of imitating or even exceeding human cognitive capacities, including sensing, language interaction, reasoning and analysis, problem solving, and even creativity."*

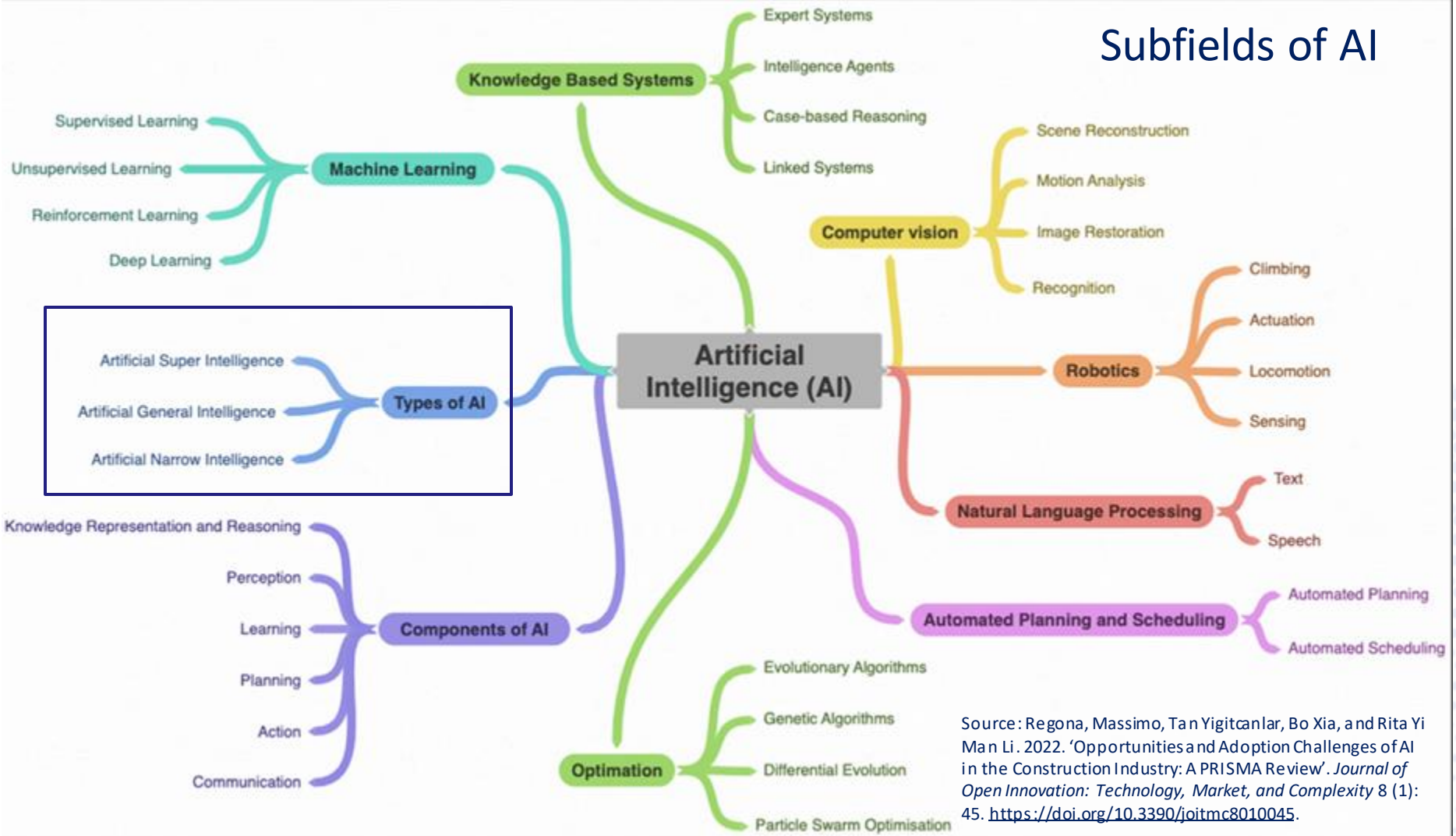
*UNESCO World Commission on the Ethics of Scientific Knowledge and Technology (2019). Preliminary Study on the Ethics of Artificial Intelligence. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000367823*

https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf




**"intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience"**

# Subfields of AI




Source: Regona, Massimo, Tan Yigitcanlar, Bo Xia, and Rita Yi Man Li. 2022. 'Opportunities and Adoption Challenges of AI in the Construction Industry: A PRISMA Review'. *Journal of Open Innovation: Technology, Market, and Complexity* 8 (1): 45. https://doi.org/10.3390/joitmc8010045.

# AI types based on capabilities

## Narrow AI

Weak AI - the only type of AI that really exists today.

- can be trained to perform a single or narrow task (even to outperform a human)
- can't perform outside of its defined task

Virtual assistants: Siri, Amazon's Alexa, IBM Watson, Chatbots: OpenAI's ChatGPT

## General AI

Artificial General Intelligence (AGI), Strong AI - a theoretical concept

If developed, it should be able to use previous learning and skills to accomplish new tasks in a different context without the need for human beings to train the underlying models. Could learn and perform any intellectual task that a human being can.

## Super AI

Super AI, artificial superintelligence and - a theoretical concept.

If ever realized, it would think, reason, learn, make judgements and have cognitive abilities that surpass those of human beings.

# Generative AI

Deep-learning models that can generate text, images and other original content types based on the data they were trained on.


‘Generative AI : A Primer’. 2023. JISC. https://beta.jisc.ac.uk/reports/generative-ai-a-primer.

## Large Language Models

- a subset of deep learning
- algorithm that can perform natural language processing tasks (recognize, translate, predict, generate text or other content
- can produce plausible but false information, often culturally or politically biased

- AI-based tools

# Computational linguistics and related areas

Using AI technologies and providing input for the development of AI tools

- CLARIN tools (discover, explore, exploit, annotate, analyse or combine language data)
- ELEXIS Pathfinder to Computational Lexicography for Developers and Computational Linguists (DARIAH)

The Open Knowledge Maps discovery tool relies on Natural Language Processing.

# Image recognition

**Project *Principal Components*: teaching algorithms to produce reusable classifications**

- Project by the National Museum of Art, Architecture and Design in Oslo
- ImageNet, written in Caffe (free license)
- Limitations: the reliability of results depends on the size and quality of the underlying image database

https://iconclassblog.com/2017/06/12/iconclass-and-ai

# Image generation

**Midjourney** https://www.midjourney.com/

- "best AI image results"
- Paid membership = own the rights
- $10/month - ~200 images/month
- can only be used through Discord

**DALL.E 3** https://labs.openai.com/

- "easy-to-use AI image generator"
- Owns input and output rights
- $15 = 115 credits - 1 credit = 1 prompt

**Getty Images** https://www.gettyimages.co.uk/ai/generation/

- "for usable, commercially safe images"
- model trained only on Getty images
- currently requires demo request

Text-to-image generation from natural language descriptions (prompts):

1. train a neural network to understand what 'things' are - destroy training data through the successive addition of random noise, and then learning to recover the data by reversing this noising process
2. Diffusion = pass randomly sampled noise through the learned denoising process and edit it in a series of steps to match interpretation of the prompt

https://zapier.com/blog/best-ai-image-generator/
https://www.assemblyai.com/blog/diffusion-models-for-machine-learning-introduction/

**Midjourney - "transparent empty trojan horse without greek soldiers inside it, cartoon style"** - Upscaled (2x)

# Image-generation

**Some serious limitations…**

prompt: “medical doctor” (because “doctor” gave some Doctor Who images)

**… trained on biased data**

# Subtitles and captions

**Whisper - Automatic speech recognition (ASR) system trained on 680,000 hours of multilingual and multitask supervised data collected from the web**

https://openai.com/research/whisper


**jargon, organisations and uppercase:**
“preprints”, “Coalition of Open Access Repositories”, “JISC”, “EOSC Synergy”, “re3data”, “PLOS Biology”, “Serpa Romeo”, “Horizon Europe”, “Creative Commons”, “Zenodo”

**BUT**
uppercase: “fair principles”, “CC by license”,
mistranscript: “coalition on us” (cOAlition S), “Yosk” (EOSC), “CC buy” (CC-BY)
names: Melissa (Milica), Antonio (Antonia)

# Chatbots

- Chatbot - software application that aims to mimic human conversation (using text or voice)
- Mostly based on Large Language Models (LLMs), i.e. language models consisting of a neural network with many parameters
- Generative AI: uses neural networks to identify the patterns and structures in existing data to generate new content
- Available to general users
- List of chatbots: https://en.wikipedia.org/wiki/List_of_chatbots

# ChatGPT

"write me a text in the style of Shakespeare explaining what Open Science is"

*Good sirs and ladies, attend ye well,*
*For I bring forth a tale of great tell,*
*Of a noble cause, both rare and fine,*
*The open science movement doth shine.*

*In days of old, knowledge was kept close,*
*Locked away in books, hidden from those,*
*Who sought the truth, the secrets of life,*
*And the answers to questions, both great and rife.*

*But now the times have changed, a new way is found,*
*Where knowledge doth flow, unbound,*
*From scientist to scientist, and far beyond,*
*The people's wisdom and knowledge is fanned.*

"explain to researchers in one small paragraph what is Open Science and why it is useful."
refined: "rewrite in plain english and shorter"

Open Science is about sharing everything in research—data, methods, and results—with everyone. It helps speed up discoveries, makes research better, and lets scientists work together more easily. By being open and transparent, scientists can avoid doing the same work twice, collaborate across different fields, and overall, make science more efficient and trustworthy. It's like teamwork for scientists and makes it easier for everyone to learn and build on each other's work.

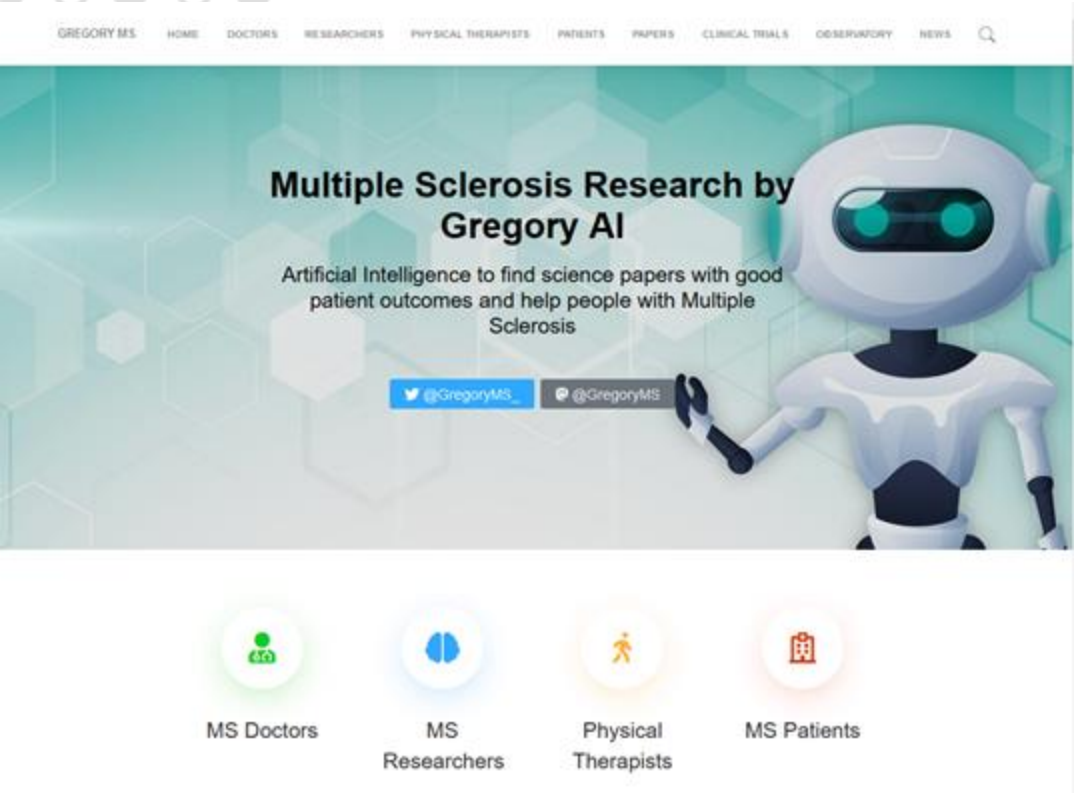
BUT careful with wrong information and fake publications

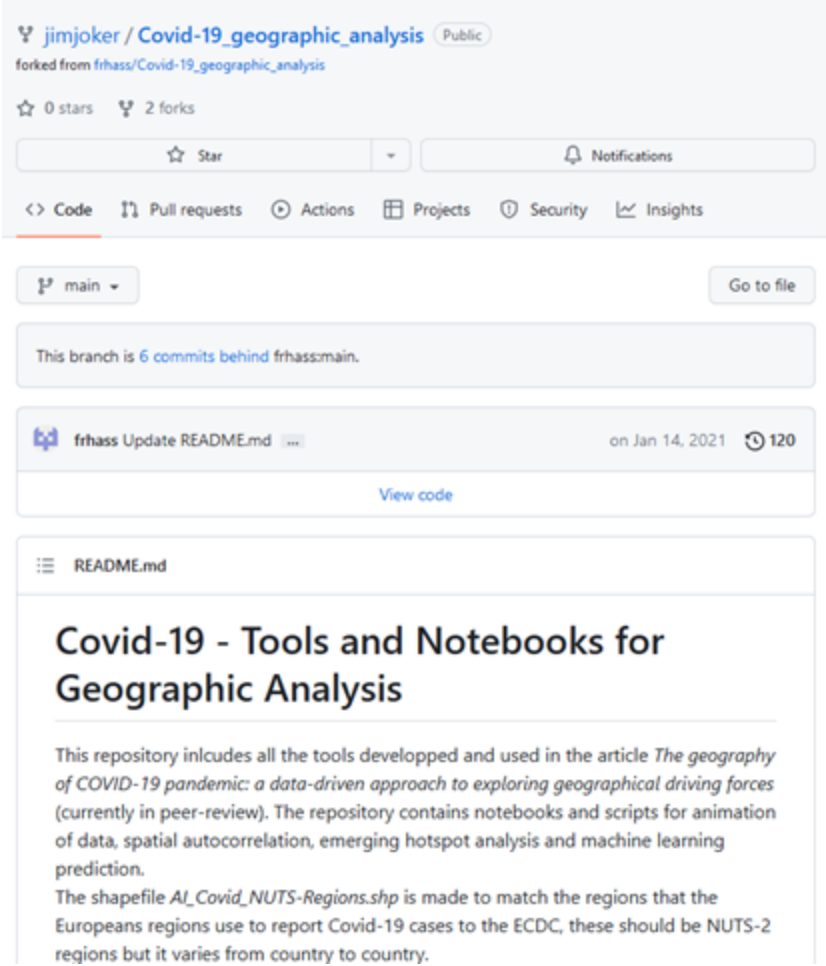
# Chatbots and academic tools

Useful account for lists of AI tools for academia: https://twitter.com/MushtaqBilalPhD

- GPT-JT - an open source alternative to chatGPT - https://huggingface.co/spaces/togethercomputer/GPT-JT
- Assistant by scite https://scite.ai/assistant - “Ask simple questions and get reliable answers from the full-texts of millions of research articles”
- Inciteful https://inciteful.xyz/ - built-in with Zotero “Build a network of academic papers and we'll analyze the network to help you discover the most relevant literature.”
- R Discovery https://discovery.researcher.life “accelerates your research discovery journey, with latest and relevant content in your area of interest.”
- PapelPal https://paperpal.com/ - academic writing
- Jenni https://jenni.ai/ - with plagiarism checked
- SciSummary https://scisummary.com/ - Use AI to summarize scientific articles in seconds
- Quillbot https://quillbot.com/ - paraphrasing, grammar check, plagiarism check

# Healthcare


GREGORY MS
HOME
DOCTORS
RESEARCHERS
PHYSICAL THERAPISTS
PATIENTS
PAPERS
CLINICAL TRIALS
OBSERVATORY
NEWS
Multiple Sclerosis Research by Gregory AI
Artificial Intelligence to find science papers with good patient outcomes and help people with Multiple Sclerosis
@GregoryMS_
@GregoryMS
MS Doctors
MS Researchers
Physical Therapists
MS Patients

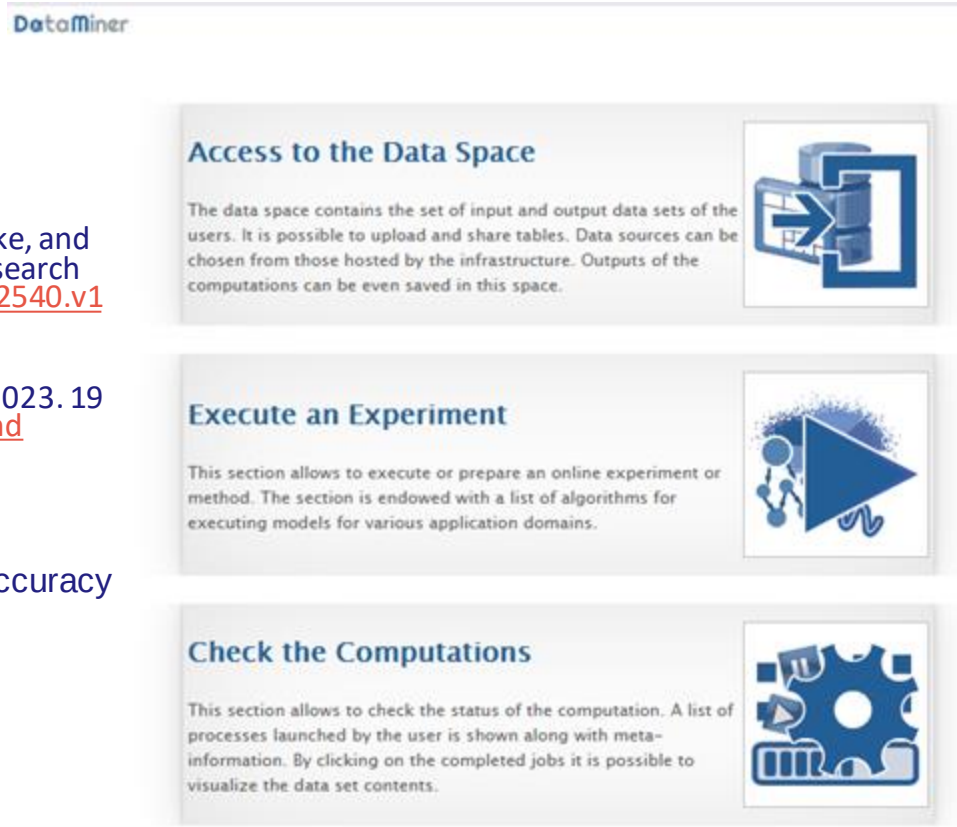

jimjoker / Covid-19_geographic_analysis
Public
forked from frhass/Covid-19_geographic_analysis
0 stars
2 forks
Star
Notifications
Code
Pull requests
Actions
Projects
Security
Insights
main
Go to file
This branch is 6 commits behind frhass:main.
frhass Update README.md
on Jan 14, 2021
120
View code
README.md
Covid-19 - Tools and Notebooks for Geographic Analysis
This repository inlcudes all the tools developped and used in the article The geography of COVID-19 pandemic: a data-driven approach to exploring geographical driving forces (currently in peer-review). The repository contains notebooks and scripts for animation of data, spatial autocorrelation, emerging hotspot analysis and machine learning prediction.
The shapefile AI_Covid_NUTS-Regions.shp is made to match the regions that the Europeans regions use to report Covid-19 cases to the ECDC, these should be NUTS-2 regions but it varies from country to country.

# Virtual labs

Klami, Arto, Theodoros Damoulas, Ola Engkvist, Patrick Rinke, and Samuel Kaski. 2022. 'Virtual Laboratories: Transforming Research with AI'. TechRxiv. https://doi.org/10.36227/techrxiv.20412540.v1

'Discover the Role of Artificial Intelligence in Virtual Labs'. 2023. 19 April 2023. https://blog.praxilabReduce the Human Error and Increase the Accuracys.com/2023/04/19/role-of-artificial-intelligence-in-virtual-labs/.

1. Reduce the Human Error and Increase the Accuracy
2. 24/7 available
3. Digital assistance
4. No putting yourself into risky situations
5. Perform repetitive jobs
6. Daily applications
7. Making decision faster



# Citizen Science and AI

- Open University, UK https://citsci.kmi.open.ac.uk/projects/
- UNDP India Leveraging Citizen Science and AI to Combat Air Pollution - model trained thanks to citizen scientists on Zooniverse
- Ponti, M., Seredko, A. Human-machine-learning integration and task allocation in citizen science. Humanit Soc Sci Commun 9, 48 (2022). https://doi.org/10.1057/s41599-022-01049-z
- Ceccaroni, L., Bibby, J., Roger, E., Flemons, P., Michael, K., Fagan, L. and Oliver, J.L., 2019. Opportunities and Risks for Citizen Science in the Age of Artificial Intelligence. Citizen Science: Theory and Practice, 4(1), p.29. DOI: http://doi.org/10.5334/cstp.241

Interesting tools:

- FASTCAT Cloud
- Pl@ntNet

# Projects

- AI4EOSC - Artificial Intelligence for the European Open Science Cloud

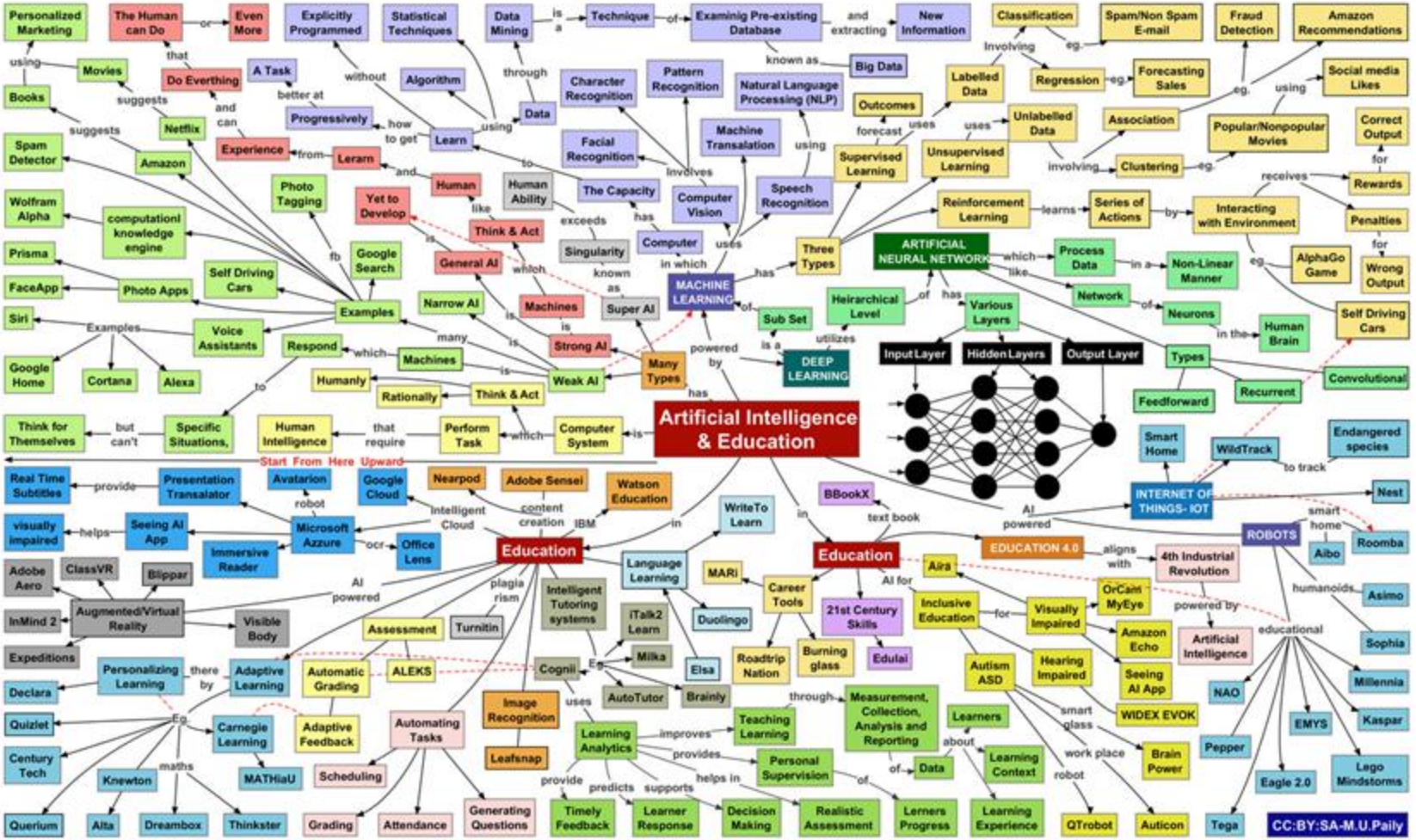
https://ai4eosc.eu/about/project-and-goals/

focus: Artificial Intelligence (AI) along with Deep Learning (DL) and Machine Learning (ML)

Increasing the service offer in the EU landscape by expanding the European Open Science Cloud (EOSC) ecosystem to support the effective utilization of state-of-the-art AI techniques by the research community

- DEEP-Hybrid- DataCloud H2020

DEEP platform, allowing researchers to exploit computing resources from pan-European e-Infrastructures.






See also: http://what-when-how.com/artificial-intelligence/artificial-intelligence-and-education/

# Challenges and recommendations

## How to Use ChatGPT to Create AI Images

ChatGPT might not be able to generate images directly, but there's a workaround for using ChatGPT to create better AI images.

BY JOHN AWA-ABUON PUBLISHED 6 DAYS AGO

HOME > TECHNOLOGY EXPLAINED

## Has ChatGPT Rendered Student Essays Obsolete?

Should students even be using ChatGPT?

BY JOWI MORALES PUBLISHED 3 DAYS AGO

## Apple becomes the latest company to ban ChatGPT for internal use

Didn't stop OpenAI rolling out iOS ChatGPT app, just made things a bit awkward

Brandon Vigliarolo Fri 19 May 2023 15:15 UTC

Digital Culture Artificial Intelligence

NEWS | 18 January 2023

## ChatGPT listed as author on research papers: many scientists disapprove

**At least four articles credit the AI tool as a co-author, as publishers scramble to regulate its use.**

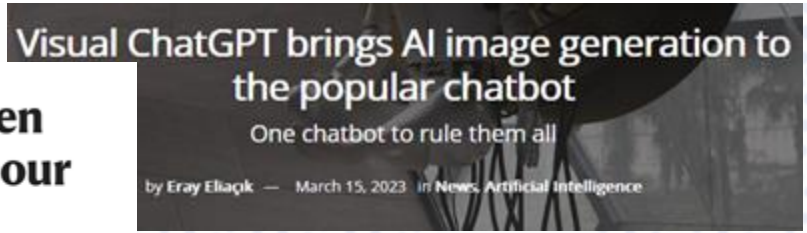
## ChatGPT essays and more: How teachers and schools are dealing with AI writing

Is artificial intelligence an academic tool for good or evil?

## 'ChatGPT-generated reading list' sparks AI peer review debate

Social scientist sees hand of ChatGPT in list of non-existent papers cited in peer reviewer's rejection

April 5, 2023




HOME > SCIENCE > VOL. 379, NO. 6630 > CHATGPT IS FUN, BUT NOT AN AUTHOR

EDITORIAL

## ChatGPT is fun, but not an author

H. HOLDEN THORP Authors Info & Affiliations

SCIENCE · 26 Jan 2023 · Vol 379, Issue 6630 · p. 313 · DOI: 10.1126/science.adg7879

10,497 60

EDITORIAL | 24 January 2023

## Tools such as ChatGPT threaten transparent science; here are our ground rules for their use

**As researchers dive into the brave new world of advanced AI chatbots, publishers need to acknowledge their legitimate uses and lay down clear guidelines to avoid abuse.**

# Concerns and criticism

- Quality and relevance data used (dirty data, relying on digital data only, poor representation in data
- Quality and relevance of the research outputs based on AI
- Bias!!!
- Copyright (EU, USA, UK)
- Combining AI techniques with poor disciplinary knowledge
- Sensationalism
- False accusations of unethical use of AI tools

An interesting case


Article | Open Access | Published: 22 September 2020

**Tracking historical changes in perceived trustworthiness in Western Europe using machine learning analyses of facial cues in paintings**

Lou Safra, Coralie Chevallier, Julie Grèzes & Nicolas Baumard

*Nature Communications* 11, Article number: 4728 (2020) | Cite this article

Discussion on PubPeer:
https://pubpeer.com/publications/57F5842590AEC0B1968B90CD18E1A0


Original Research | Open Access | Published: 12 January 2021

**A set of distinct facial traits learned by machines is not predictive of appearance bias in the wild**

Ryan Steed & Aylin Caliskan

*AI and Ethics* 1, 249–260 (2021) | Cite this article

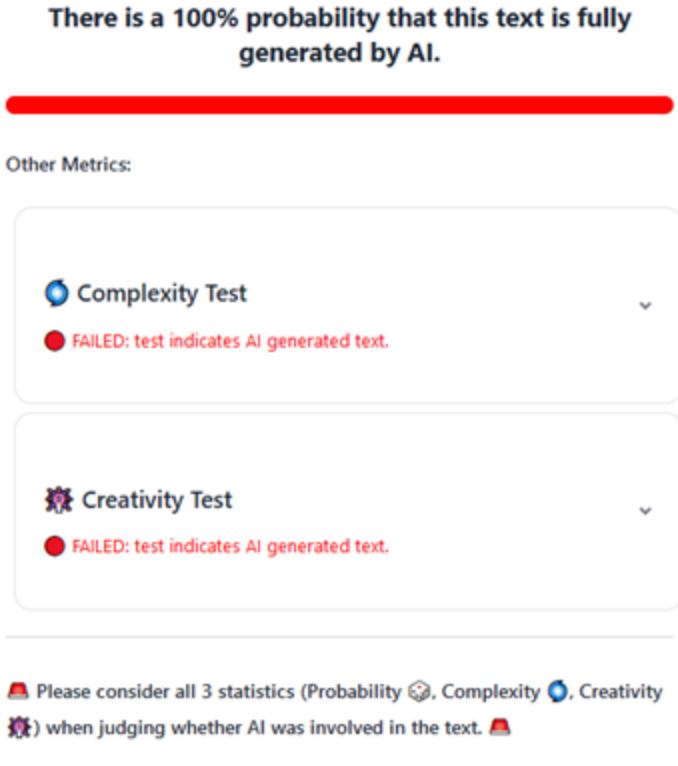

**How AI is hijacking art history**

Published: November 1, 2021 1.24pm CET

Sonja Drimmer
Associate Professor of Medieval Art, UMass Amherst

# Detecting text generated by chatbots

- GPT4 Detector .ai
- AI Text Classifier
- GPTZero
- Winston AI

Technology

**Plagiarism tool gets a ChatGPT detector – some schools don't want it**

Popular plagiarism detection software used by many schools and universities worldwide is set to get an AI-detecting component in the wake of the release of ChatGPT

By Jeremy Hsu

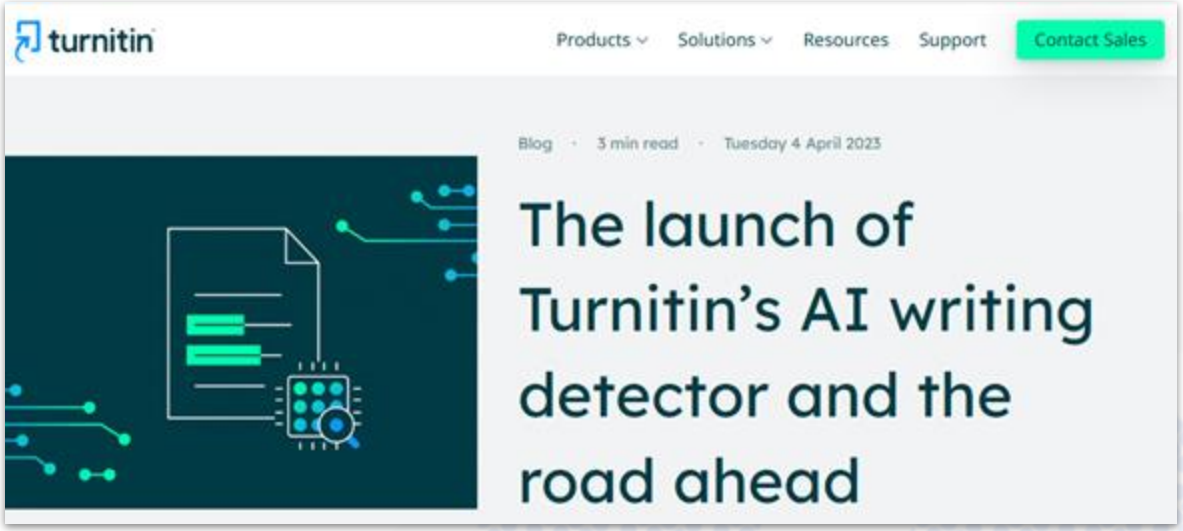
3 April 2023

TECH IN YOUR LIFE

**We tested a new ChatGPT-detector for teachers. It flagged an innocent student.**

Five high school students helped our tech columnist test a ChatGPT detector coming from Turnitin to 2.1 million teachers. It missed enough to get someone in trouble.

Analysis by Geoffrey A. Fowler



turnitin

Products Solutions Resources Support Contact Sales

Blog · 3 min read · Tuesday 4 April 2023

**The launch of Turnitin's AI writing detector and the road ahead**

FUTURE

**How to Prove You Didn't Use ChatGPT: One Simple Trick to Avoid ChatGPT Plagiarism Accusations**

By Amy D

Posted on May 22, 2023

"Today, we are pleased to announce the launch of our AI writing detection capabilities in Turnitin Feedback Studio (TFS), TFS with Originality, Turnitin Originality, Turnitin Similarity, Simcheck, Originality Check, and Originality Check+. The detector will support over 2.1 million educators and more than 10,700 institutions, reaching more than 62 million students. It is a milestone that represents an incredible commitment from the Turnitin team as well as our customers."

https://www.turnitin.com/blog/the-launch-of-turnitins-ai-writing-detector-and-the-road-ahead

# Chatbots, ChatGPT, and Scholarly Manuscripts

## WAME Recommendations on ChatGPT and Chatbots in Relation to Scholarly Publications

1. Only humans can be authors;
2. Authors should acknowledge the sources of their materials;
3. Authors must take public responsibility for their work;
4. Editors and reviewers should specify, to authors and each other, any use of chatbots in evaluation of the manuscript and generation of reviews and correspondence; and
5. Editors need appropriate digital tools to deal with the effects of chatbots on publishing.

In addition, this revision acknowledges that chatbots are used to perform different functions in scholarly publications. Currently, individuals in scholarly publishing may use chatbots for: 1) simple word-processing tasks (analogous to, and an extension of, word-processing and grammar-checking software), 2) the generation of ideas and text, and 3) substantive research. The Recommendations have been tailored for application to these different uses.

https://wame.org/page3.php?id=106

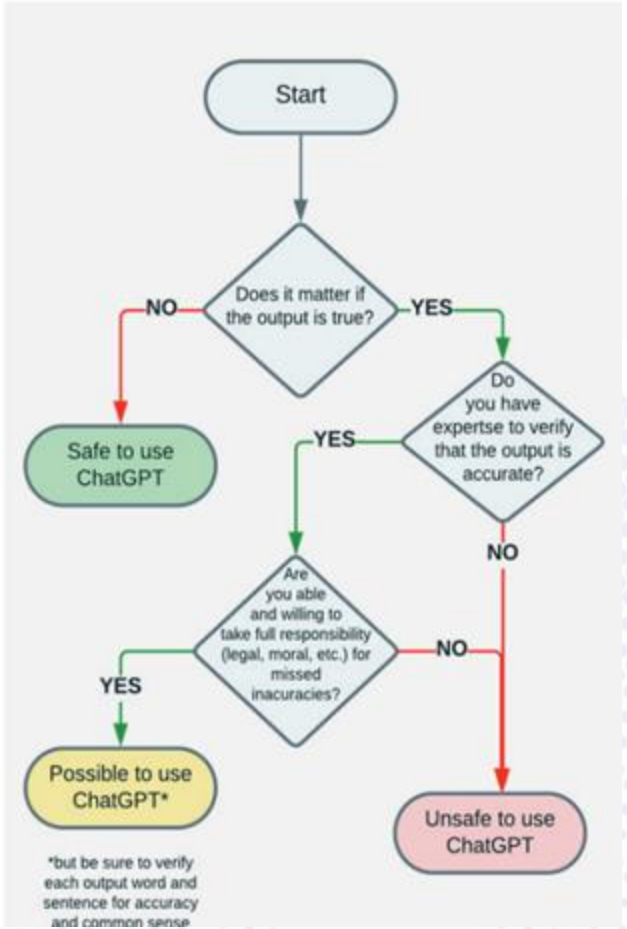
# Ethics


- UNESCO. 2023. ‘UNESCO’s Recommendation on the Ethics of Artificial Intelligence: Key Facts’. https://unesdoc.unesco.org/ark:/48223/pf0000385082.
- ‘WHO Calls for Safe and Ethical AI for Health’. n.d. Accessed 28 November 2023. https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health.
- Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. Zenodo. https://doi.org/10.5281/zenodo.3240529
- Responsible AI licences (open RAIL): https://www.licenses.ai/ai-licenses


# Safe use of AI in research and education

- Sabzalieva, Emma, and Arianna Valentini. 2023. ‘ChatGPT and Artificial Intelligence in Higher Education: Quick Start Guide - UNESCO Digital Library’. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385146.
- And more from UNESCO: https://www.unesco.org/en/digital-education/artificial-intelligence
- ‘Generative AI : A Primer’. 2023. JISC. https://beta.jisc.ac.uk/reports/generative-ai-a-primer.

Figure 1: When is it safe to use ChatGPT?[5]



Global Grand Challenges

BILL & MELINDA GATES foundation

ABOUT PARTNERSHIPS CHALLENGES AWARDED GRANTS GRANT OPPORTUNITIES NEWS

## Catalyzing Equitable Artificial Intelligence (AI) Use

SHARE THIS

APPLY FOR THIS OPPORTUNITY →

**INITIATIVE**
Grand Challenges

**DATE OPEN**
May 22, 2023, 4:00 am PDT

**DEADLINE**
Jun 05, 2023, 11:30 am PDT

**SUPPORTING MATERIALS**
- Request for Proposal
- Rules and Guidelines
- Application Instructions

https://gcgh.grandchallenges.org/challenge/catalyzing-equitable-artificial-intelligence-ai-use

# Relevance for OS trainers

- An emerging training topic.
- Some OS projects develop AI tools (and some also use open data to train algorithms)
- AI revolves around data (quality, curation, storage…)
- AI tools and environments involved data discovery and RDM
- Educational tools based on AI
- RRI context - privacy, bias, plagiarism, ethical issues
- Being familiar with best practice / recommendation
- Everybody's talking about it.


OECD. 2023. Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research. OECD. https://doi.org/10.1787/a8d820bd-en

# Training challenges

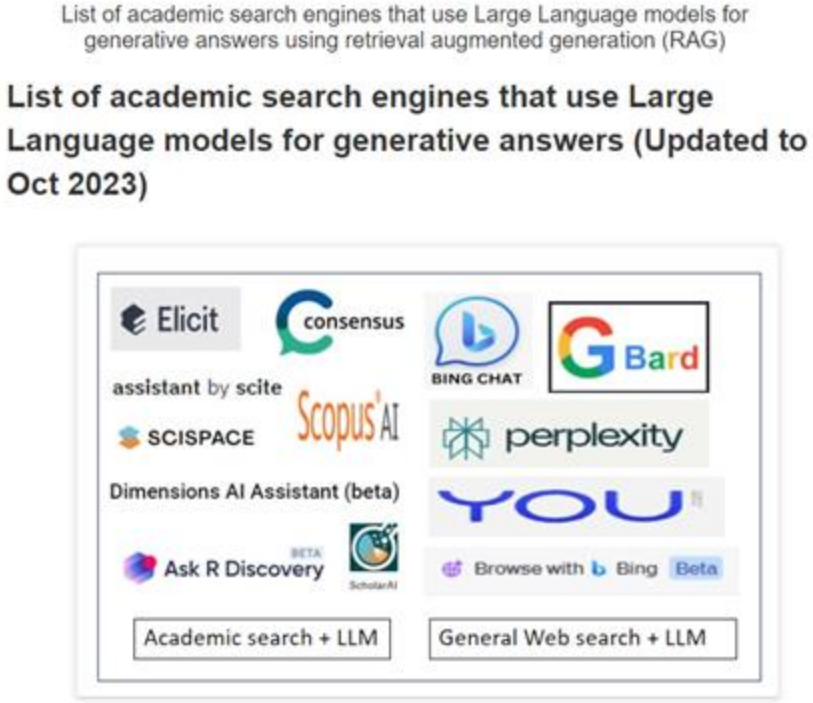
- Wide and unclear scope (what AI is, what AI isn't, and what isn't AI)
- Technical knowledge/understanding is required
- Diverse perspectives and trainees' perceptions
- Explaining to researchers that data quality matters
- Sensationalism in the media
- Popular perceptions

Useful readings: https://www.zotero.org/groups/5131206/generative_ai_and_related_phenomena

# Training tips

- Be informed and try to understand the operation and limitations of AI tools
- Test, demonstrate, analyze
- Avoid sensationalism
- Highlight challenges and open issues
- Focus on information literacy

List of academic search engines that use Large Language models for generative answers using retrieval augmented generation (RAG)




Tay, Aaron. n.d. 'List of Academic Search Engines That Use Large Language Models for Generative Answers Using Retrieval Augmented Generation (RAG)'. Aaron Tay's Musings about Librarianship. Accessed 16 November 2023. https://musingsaboutlibrarianship.blogspot.com/p/list-of-academic-search-engines-that.html.

# THANKS

**Web**
www.openaire.eu

**Email**
milica.sevkusic@eifl.net
jonathan.england@openaire.eu

**Twitter**
@openaire_eu
@lessormore4
@jonatortue