Imperial College London

The role of artificial intelligence tools in climate change scientific assessments

Overview of AI approaches and tools: AI landscape and tools for scientific assessments
USGCRP Advisory Committee's Fall 2024 meeting

Alaa Al Khourdajie
Research Fellow, Chemical Engineering, Imperial College London
25/09/2024

Agenda

1 IPCC Mandate

2 Challenges of undertaking scientific assessments

3 Leveraging AI-driven solutions for scientific synthesis

4 Communicating scientific assessments' findings using generative AI

IPCC Mandate

IPCC Mandate

"… to <u>**assess**</u> **on a comprehensive, objective, open and transparent basis** the scientific [literature].."

"IPCC reports should be **neutral with respect to policy**…"

"The assessments are **policy-relevant but not policy prescriptive**…"

Source:

- *Principles Governing IPCC Work*
- *IPCC Factsheet*

IPCC Mandate

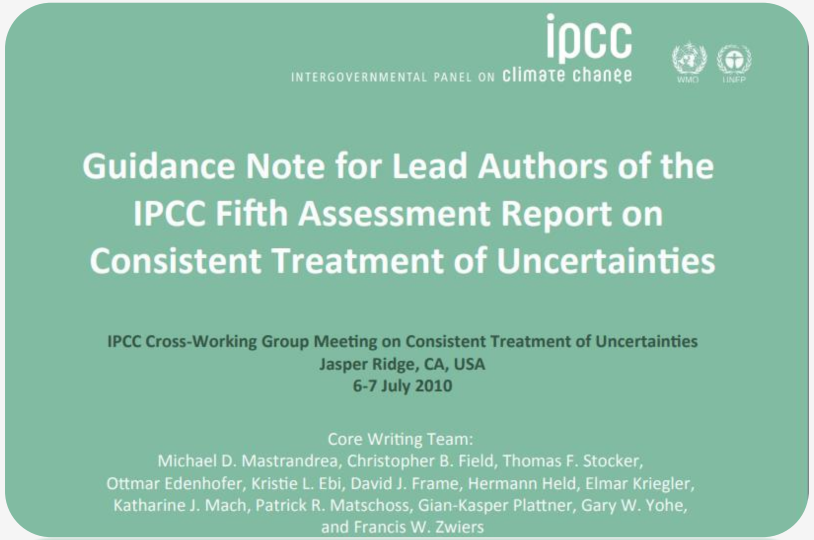
"… to **assess**…"

Identifying and communicating areas of **scientific consensus**

Several **methods** for undertaking **scientific synthesis**:

- such as meta-analyses, systematic maps and reviews, and weight of evidence
- to maximise rigour and minimise susceptibility to bias and errors

Treatment of uncertainties



"… to **assess**…"

- Calibrated uncertainty language to express **confidence levels** about **consensus**
- Confidence expressed using five **qualifiers**: ***very low, low, medium, high and very high***.

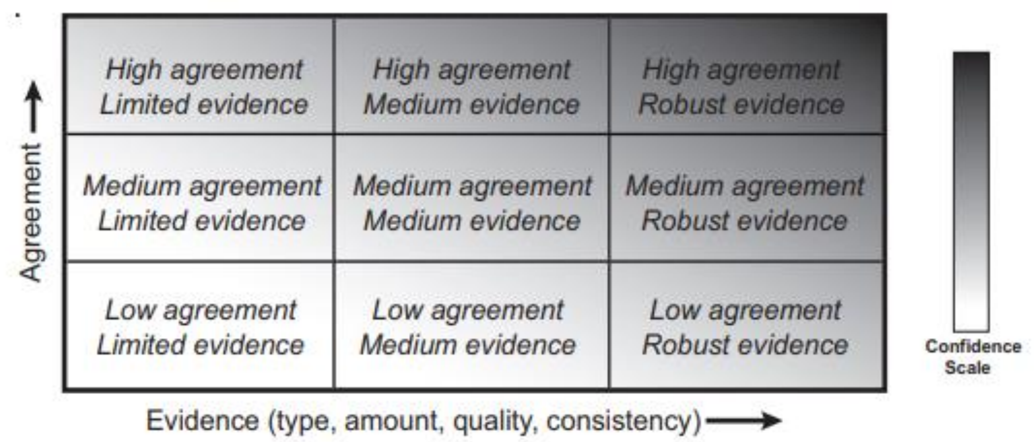


Figure 1: A depiction of evidence and agreement statements and their relationship to confidence. Confidence increases towards the top-right corner as suggested by the increasing strength of shading. Generally, evidence is most robust when there are multiple, consistent independent lines of high-quality evidence.

Challenges of undertaking scientific assessments

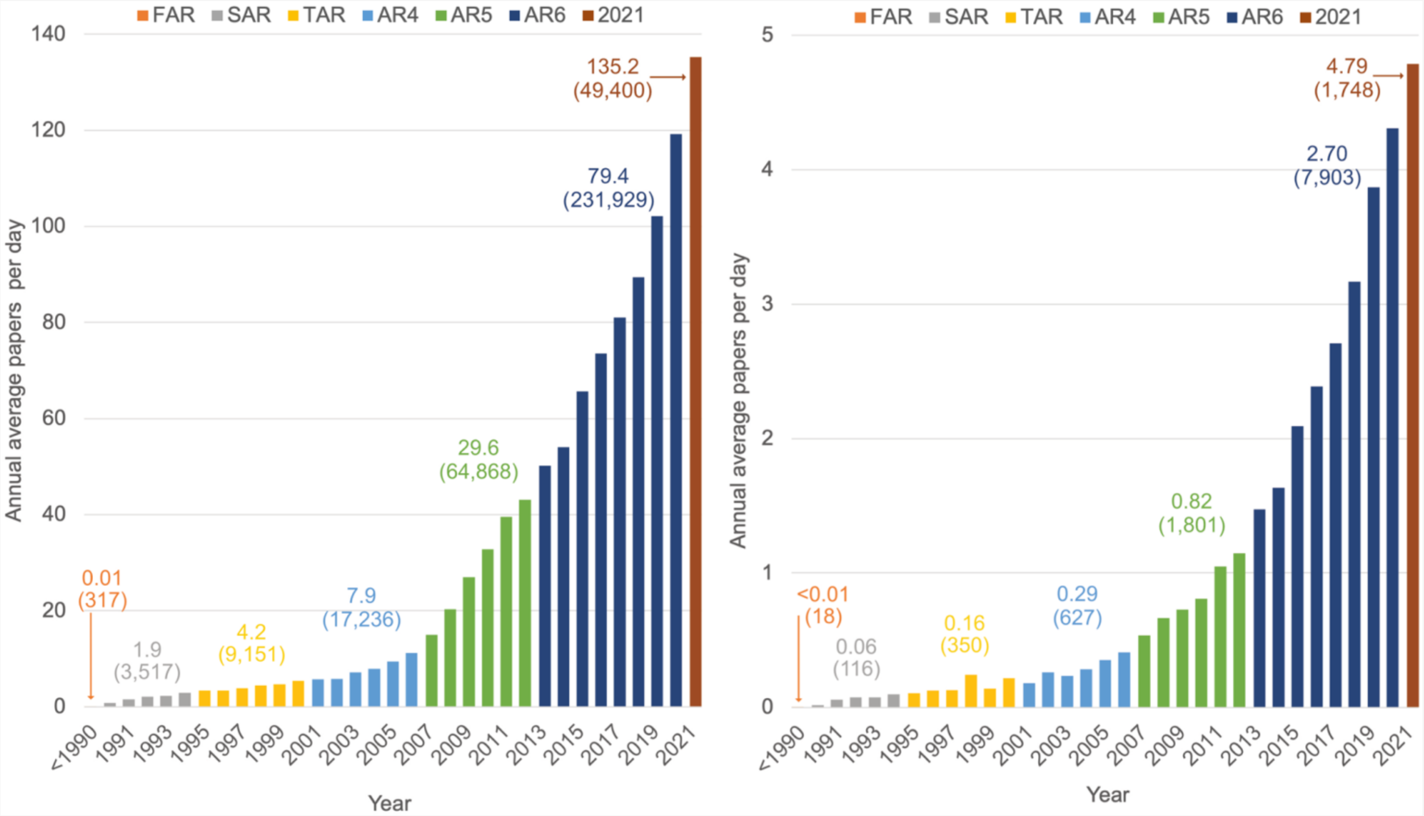
Key challenges

1)

The **exponential growth and increasing complexity of scientific literature** complicating the IPCC's mandate to conduct '***comprehensive, objective, open and transparent***' assessment in line with the IPCC principles

Not to mention the workload required from scientists to engage with assessment

Growing number of publications [daily average vs (cumulative)]



"climate change" or "global warming" (left),
"extreme rainfall" or "heavy precipitation" (right)

De-Gol et al., 2023

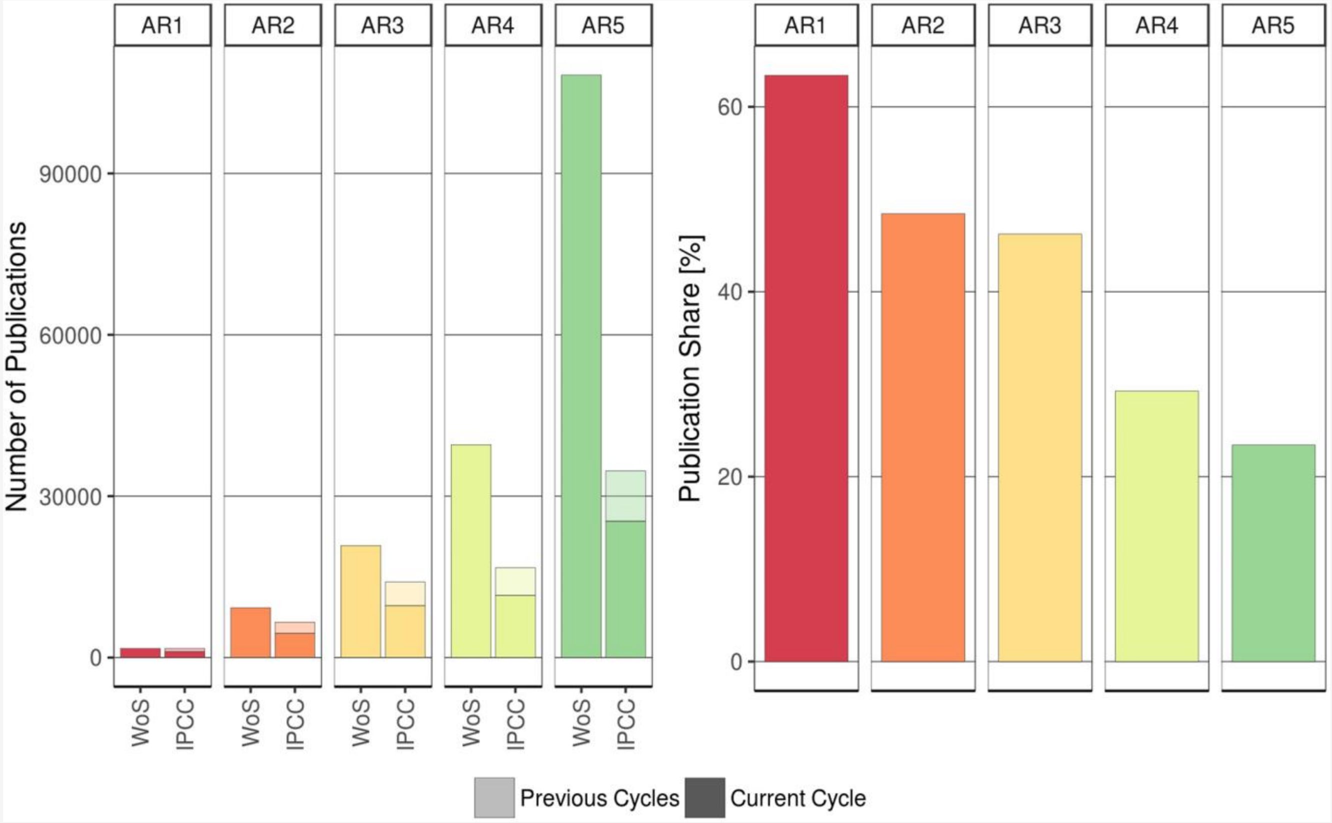
Key challenges

1)

The **exponential growth and increasing complexity of scientific literature** complicating the IPCC's mandate to conduct '***comprehensive, objective, open and transparent***' assessment in line with the IPCC principles

Not to mention the workload required from scientists to engage with assessment

The number of references in each assessment report has been declining rapidly as a proportion of relevant available literature over time



Left: Total number of scientific articles published (WoS) and total number of unique references (IPCC). Right: "Publication share is the ratio"

Minx et al., 2017

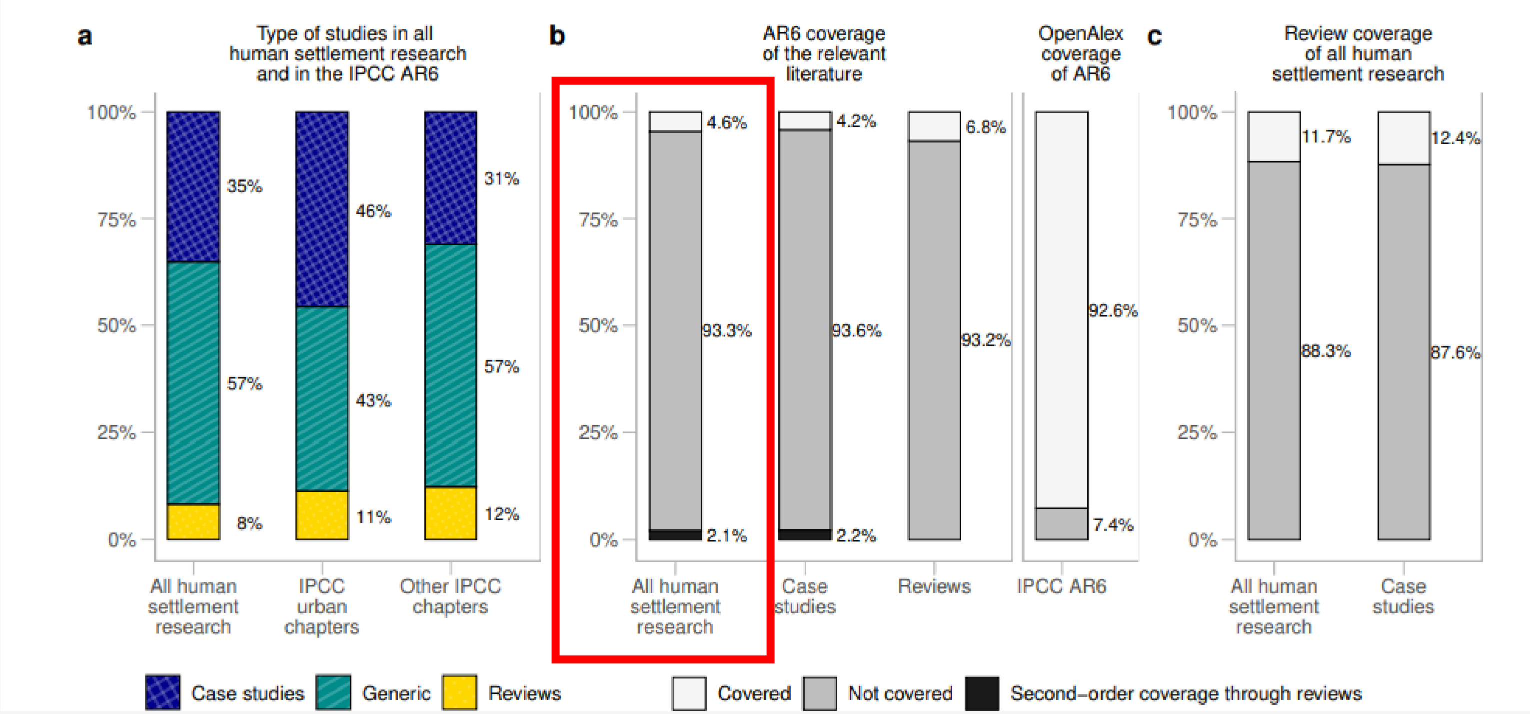
Key challenges

1)

The **exponential growth and increasing complexity of scientific literature** complicating the IPCC's mandate to conduct '***comprehensive, objective, open and transparent***' assessment in line with the IPCC principles

Not to mention the workload required from scientists to engage with assessment

IPCC AR6 covered only 4.6% of research related to human settlement



Type of references in all human settlement research and the IPCC and the share of the covered literature

Montfort et al, 2024

Key challenges

2)

The **length of the IPCC reports**

AR6 produced a total of ~12,000, including WGs, SRs and ~MR

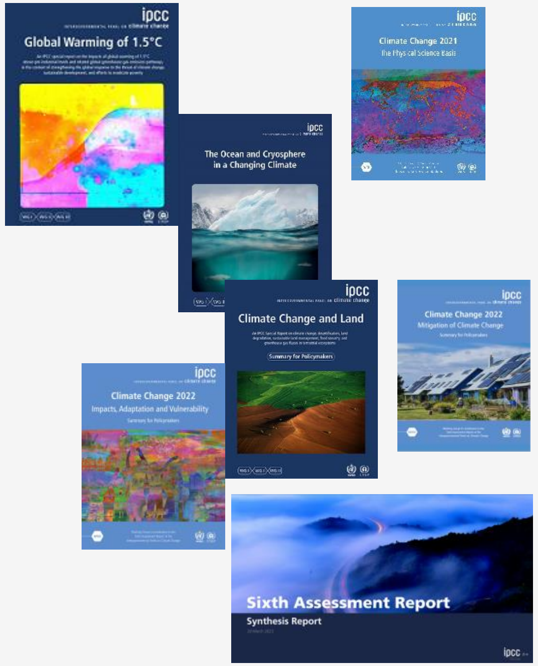
Challenges for:

- **accessibility of the detailed underlying evidence** in the main reports (despite the SPMs and TSs)
- dissemination of **important scientific findings that do not feature in the summaries**

Key challenges

1)

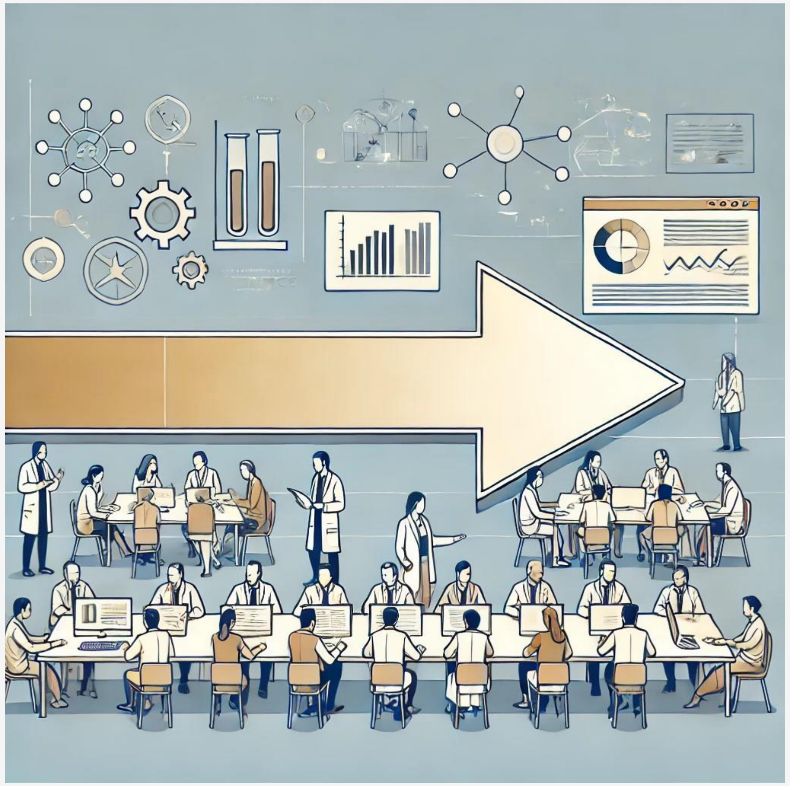
The **exponential growth and increasing complexity of scientific literature**



© DALL.E

2)

Dissemination of **important scientific findings**



© DALL.E

Leveraging AI-driven solutions for scientific synthesis

Solutions: NOT ChatGPT

Early Warning and Assessment Division

UN environment programme

4 December 2023

APPROVED DOCUMENT

Disclosure on the use of Generative Artificial Intelligence-based tools in UNEP's seventh edition of the Global Environment Outlook (GEO-7)

APPROVED DOCUMENT

4 December 2023

Guidance from the Global Environment Outlook's Multidisciplinary Expert Scientific Advisory Group (MESAG) on the use of Generative Artificial Intelligence (AI)-based tools in the context of UNEP's seventh edition of the Global Environment Outlook (GEO-7)

Solutions: Systematic literature review

- **Systematic or evidence maps:**
  - A type of systematic literature review
  - Catalogues available evidence on a broad topic
  - Aim to characterise the overall landscape of research in a field

- **AI-enhanced approach**:
  - Using AI to scale up the traditional systematic map methodology
  - Allows to handle a much larger volume of literature than would be possible with manual methods

Solutions: Overview

A set AI/ML methods aim to process and synthesise large volumes of scientific literature

Table 1: Approaches for scientific assessments

Approaches and tools	Method	Output	Strengths	Limitations	Examples
Systematic synthesis and evidence maps	LLMs and ML to identify and classify articles, and draw evidence maps	Systematic evidence maps and interactive study database	Rapid and comprehensive coverage, identification of knowledge gaps	Limited analytical depth	Callaghan set al. (2021)
Topic modelling	Unsupervised ML for analysing word co-occurrences to identify themes	Topographic map of research publications, and clustering of topics	Reveals level of representation and gaps in literature, and stimulates future work	Limited analytical depth	Callaghan et al. (2020), Lamb et al. (2019)
Double-stacked expert-ML research approach	ML and experts' evaluation for comprehensive typology and thematic mapping of literature	Thematic clusters, and content-based insights	Leverages ML efficiency with essential expert analysis	Significant expert input	Creutzig et al. (2021)
Collaborative technology platforms (ScienceBrief.org)	Asynchronous and connected platforms for augmenting expert collaboration	Scientific briefs, and emerging scientific findings for media coverage	Keeping assessments up-to-date	Low expert engagement incentives, and no automatic identification of updated literature	ScienceBrief.org (2017-2021) De-Gol et al. (2023)
Literature exploration tools (LitMaps, SciSpace, Elicit, Scopus AI, Sematic Scholar, Google Gemini)	Network analysis and visualisation, and NLP technology for navigating and summarising scientific literature	Research article networks, summarisations, extracting data, and concept maps	Swift navigation and discovery, and simplifies subsequent literature synthesis	Accuracy and reliability concerns, and coverage is subject to the used databases	LitMaps (2024), SciSpace (2024), Elicit (2024), Scopus AI (2024), Google Deepmind (2024)

Solutions: Common elements and characteristics

- **Use of AI/ML**: All approaches leverage AI and machine learning to handle large datasets
  - Natural language processing (NLP), unsupervised ML techniques
  - Following specific search criteria (keyword-based search query) and machine learning classifier
  - Words co-occurrences, e.g. for clustering

- **Scalability**: Designed to process much larger volumes of literature than traditional manual methods

- **Automation**: Reduce manual effort in identifying, classifying, and synthesising the literature

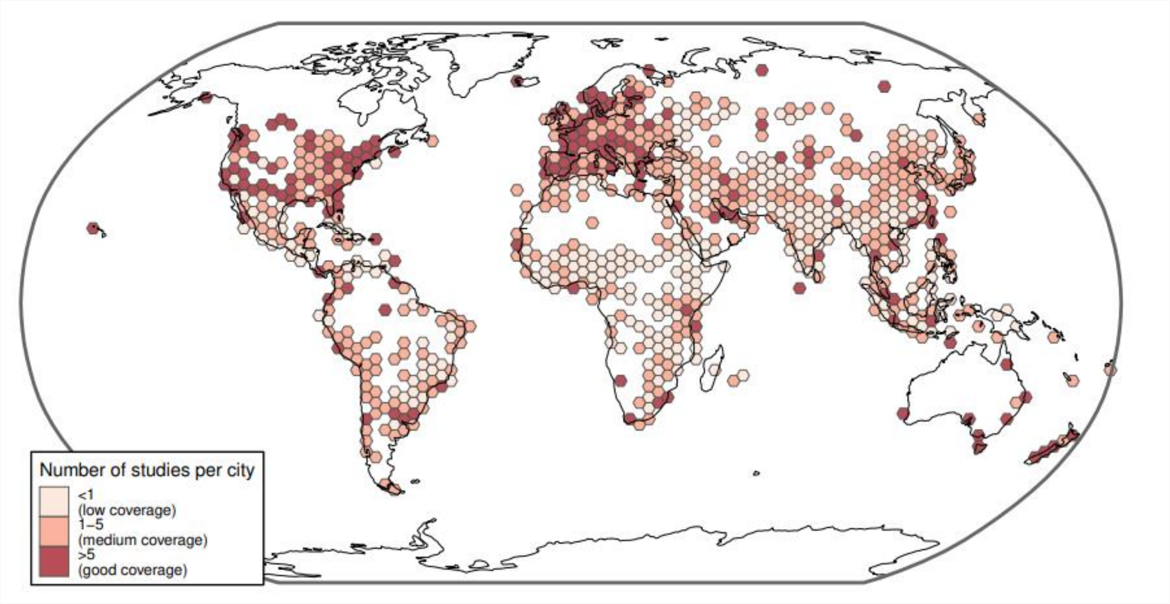
Solutions: Workflow

- **Data collection**: Gather large volumes of literature from databases (WoS, OpenAlex, …)
- **Pre-processing**: Clean and prepare text data for analysis
- **AI/ML Analysis**: Apply various techniques (NLP, topic modeling, classification) to process the literature
- **Degree of expert involvement**: Ranges from minimal (fully automated) to significant (expert-guided) in order to refine the results or provide additional analyses
- **Analytical depth**: Trade-off between breadth of coverage (*fully automated*) and depth of analysis (*expert-guided*)
- **Potential output**: ….

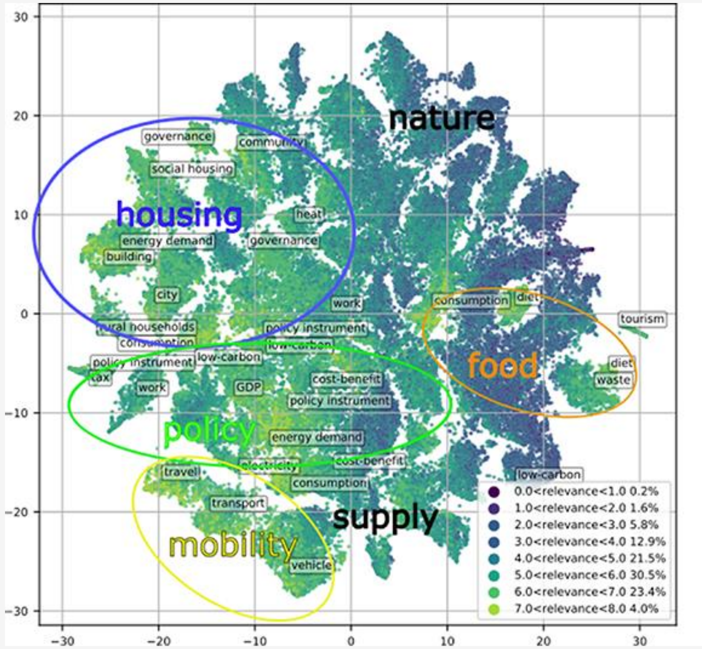
Solutions: Workflow

- **Potential output**: Varies from **evidence or topographic maps** to **thematic clusters** to **content-based insights** to **living databases**

- **Visualisation**: of maps and clusters to aid understanding



https://cc-cities.adibilis.ch/ Montfort et al, 2024



Example on demand-side mitigation by Creutzig et al. (2021)

Solutions: Issues and potential biases

- Incomplete coverage of topics based on the databases used, geographical scopes, and cultures
  - Arguably, we suffer from this under current practices too

- Limitation to English-language literature (goes against the rising importance of involving indigenous knowledge and local knowledge)

- Experts' engagement require further incentives

- Equity considerations need to be at the forefront:

  - Important to avoid discouraging scientists not engaged with these tools from participating
  - These tools require internet access, which may not be readily or continuously available to scientists in the Global South

Questions and next steps

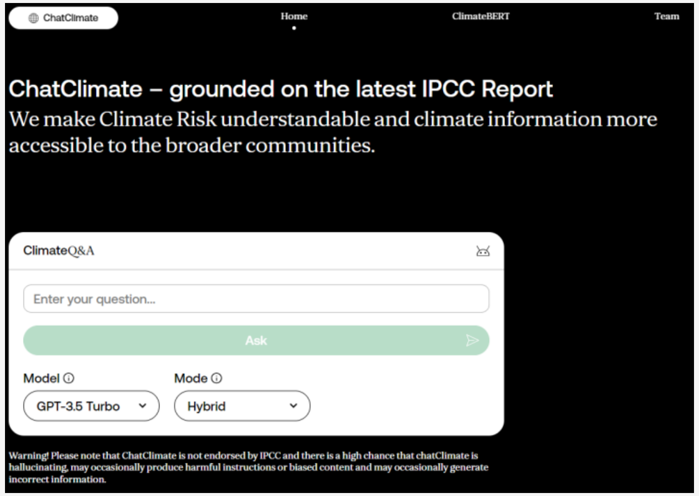
- A mapping of **different AI/ML tools** and methods that can assist in the assessment of **different types of literature** and **stages of the assessment process**

- Whether it is necessary to **update the guiding principles** of assessments processes to preserve the integrity of scientific assessment, and how?

- What is the ideal workflow for an **expert-led, AI-supported assessment** to preserve the integrity of scientific assessment?

Communicating scientific assessments' findings using generative AI

AR6 in ChatBots: ChatClimate



- Three setups (AR6 standalone, GPT-4, and Hybrid)
  - Hybrid Model: most accurate
  - Hallucination : Best managed by AR6 standalone model

- Address hallucination: bespoke prompt engineering, and top ranked tokens retrieval

Q2) When will we reach 1.5 °C?	AR6 standalone	GPT-4 knowledge standalone
Hybrid We are likely to reach 1.5°C of global warming be- tween 2030 and 2052 if the current rate of warming continues (high confidence) (IPCC AR6, Reference: IPCC AR6 WGI Chapter01, Page: 44). However, the exact timing may vary depending on the emissions re- duction pathways and efforts to achieve net-zero CO_2 emissions (In-house knowledge).	We are likely to reach 1.5 °C of global warming between 2030 and 2052 if it continues to increase at the current rate (high confidence). Reference: IPCC AR6 WGI Chapter01, Page: 44.	As an AI, I cannot predict the exact year when global temperatures will reach 1.5 °C above pre-industrial levels. However, according to the Intergovernmental Panel on Climate Change (IPCC), we could reach this threshold between 2030 and 2052 if global warming continues at its current rate. The actual timeline depends on various factors, including greenhouse gas emissions, climate policies, and global efforts to mitigate climate change.
accuracy: 5	**accuracy: 4**	**accuracy: 2**

Concerns about risks and issues reported in the lit

- Grasp of scientific concepts:
  - LLMs do not understand language
  - They are pattern recognisers, predicting text based on trained data
  - Predicting what comes next in a sequence of words
  - The pattern-spotting mechanism might result in different sequences

- Should not **re-**synthesis

"a feature, not a bug"

- Lack vital context needed to avoid misinformation

- Risks of misinterpretation, so called 'hallucinations'

- Potential biases that could influence policy advice

- Ethical challenges related to transparency and reproducibility, hindering the verification and reproduction of results

LLMs: Minimising hallucination

- **During the initial stages of development**: training of the model to capture experts' preferences and rankings (e.g. reinforcement learning by human feedback), **Alignment**

- **After training the model**: adapt and optimise its performance (e.g. additional training data – AR6 reports, and limiting the choice of returned text to top-ranked tokens – by tweaking its hyperparameters), **Fine-tuning**

- **Divides the process into two parts**: retrieval & generation models - Retrieval-augmented generation (**RAG**)

- **Prompt engineering**: instructions given to a model on how to process and respond to a user's query

Potential solutions to explore

- LLMs as "**Smart search tools**", with zero tolerance for **re**-synthesising, exclusively from the reports

- **Retrieval Hierarchy** (top ranked tokens retrieval)
  - SPMs => TSs => Chapters' ESs => Chapters' Text (for additional detail)

- Returning **FULL** paragraphs, verbatim (chunking process)

- Customised LLMs design, deeply integrated in the drafting and review process
  - tokenisation, embedding, chunking processes, prompt engineering, Fine-tuning and automated fact checking methods

Imperial College London

Thank you

QR code to the pre-print

Al Khourdajie (2024). AI & Assessments
25/09/2024