Artificial Intelligence (AI) Ethics Recommendations for the Geoscience Community

Task Group on Artificial Intelligence (AI) in Geosciences of the Commission on Geoethics of the International Union of Geological Sciences (IUGS)

26th September 2025

[1]Paul H Cleverley, [2]Mrinalini Kochupillai, [3]Mark Lindsay and [4]Emma Ruttkamp-Bloem

[1] Robert Gordon University, Aberdeen, Scotland, UK, [2] Hanken School of Economics, Helsinki, Finland, [3] The Commonwealth Scientific and Industrial Research Organisation and the University of Western Australia, [4] University of Pretoria, South Africa, Chair of UNESCO's World Commission on the Ethics of Scientific Knowledge, and Technology (COMEST), former member of the UN Secretary General's High Level Advisory Body on AI

Commission on Geoethics

IUGS

International Union of Geological Sciences

International Association for Promoting Geoethics

Reviewed By:

NAME	AFFILIATION	COUNTRY
Jonathan D. Arthur	Executive Director, American Geosciences Institute (AGI)	USA
Natalie Brand	Geological Society of South Africa (GSSA)	South Africa
Mariana de Brito	Helmholtz Center for Environmental Research (UFZ)	Germany
Simon Burnett	Information Management in the School of Law and Social Sciences, Robert Gordon University	Scotland, United Kingdom
Philippe Courtial	Executive Secretary, European Geosciences Union (EGU)	France
Giuseppe Di Capua	Istituto Nazionale di Geofisica e Vulcanologia (INGV)	Italy
Patrick Friend	Global Strategic Advisor, Marine, Coastal and Environmental Sciences, Integral Consulting Inc.	USA
Christopher Keane	Director Geoscience Profession and Higher Education, American Geosciences Institute (AGI)	USA
Andrew Kingdon	Head of Artificial Intelligence and Informatics Research, British Geological Survey (BGS)	United Kingdom
Jens Klump	Senior Principal Researcher, The Commonwealth Scientific and Industrial Research Organisation (CSIRO) Mineral Resources	Australia
Lawrence Omari-Mensah	Executive Secretary, West African Institute of Mining, Metallurgy and Petroleum (WAIMM)	Ghana
Stalyn Paucar	Universidad Central del Ecuador	Ecuador
Dean Pereira de Melo	Federal University of Minas Gerais	Brazil
Peter van der Beek	President, European Geosciences Union (EGU). Institut für Geowissenschaften, Universität Potsdam	Germany
Billy Williams	Executive Vice President, American Geophysical Union (AGU)	USA
Steven Zimmerman	Austria Computing Society	Austria

Approved By the Governing Board of the IUGS Commission on Geoethics on 30 September 2025:

NAME	AFFILIATION	COUNTRY
Silvia Peppoloni	Chair, IUGS Commission on Geoethics; Istituto Nazionale di Geofisica e Vulcanologia (INGV)	Italy
Peter T. Bobrowsky	Vice-Chair, IUGS Commission on Geoethics; Geological Survey of Canada; International Association for Promoting Geoethics (IAPG)	Canada
Giuseppe Di Capua	Treasurer, IUGS Commission on Geoethics; Istituto Nazionale di Geofisica e Vulcanologia; International Association for Promoting Geoethics (IAPG)	Italy
Ruth Allington	Councilor, IUGS Commission on Geoethics; International Association for Promoting Geoethics (IAPG)	United Kingdom
Soumaya Ayadi-Maasri	Saint Joseph University, Beirut; International Association for Promoting Geoethics (IAPG)	Lebanon
Roberto Lencina	Universidad Nacional de Tucumán; International Association for Promoting Geoethics (IAPG)	Argentina
David Mogk	Montana State University; International Association for Promoting Geoethics (IAPG)	USA

EXECUTIVE SUMMARY

This report presents recommendations for the ethical application of Artificial Intelligence (AI) within the geosciences. It has been prepared by the Task Group on AI Ethics in Geosciences under the Commission on Geoethics of the International Union of Geological Sciences (IUGS).

Statistical and machine learning approaches have supported geosciences for decades and the rapid adoption of frontier and generative AI have introduced new risks and amplifies existing ones. The study has been commissioned reflecting the rapidly increasing development and take-up of frontier generative AI solutions by geoscientists, wider use of AI in all aspects of the geosciences, and numerous high profile ethical issues in geoscience AI deployments.

This report is intended to guide practicing academic, industry, governmental and non-governmental geoscientists, society leaders, and policymakers. It outlines ethical considerations for the use of AI in geoscience, while also contributing to ongoing academic discussion on AI and geoethics.

Eight themes were identified with recommendations intended to address ethical issues and concerns occurring, or likely to occur, within the geosciences. These are: *Use AI Responsibly*; *Promote Transparency and Explainability*; *Consider Bias and Fairness*; *Obtain Informed Consent*, *Protect Personal Data*; *Practice Participatory Design and Community Engagement*; *Advocate for Environmental Protection*; *Integrity in Science, Publishing and Education*; and *Consider Geopolitics*. A simple practical one-page summary is presented for AI Ethics in the Geosciences, supported by a rich discussion of each theme with illustrative geoscience specific examples and reflective questions.

The use of virtue ethics emphasises the approach: that ethics is not just about rules and consequences, it is about moral virtues like honesty, courage and empathy, and intellectual virtues such as practical wisdom and intuition. At its core, virtue ethics asks: *What would a good and wise geoscientist do?*

A high-level roadmap for continuous improvement is proposed to keep pace with this rapidly evolving space, consisting of practical and easy to use ethical impact and risk assessments. The report also proposes establishing a standing international body within the IUGS Commission on Geoethics to advise on AI ethics and remain agile to developments and emerging new issues. The recommendations aim to foster a responsible, just, and sustainable integration of AI that serves the public good, upholds human rights, and contributes meaningfully to the integrity of the geoscience discipline and stewardship of the earth.

INTRODUCTION

Artificial Intelligence (AI) allows machines to mimic human intelligence[1] often using data whose scale exceeds what humans can analyse. AI, therefore, presents a range of opportunities for the geosciences to improve productivity, reduce uncertainty in models and stimulate new knowledge discovery. AI can support the geosciences in contributing to the UN Sustainable Development Goals (SDGs)[2]. For example, AI can help predict the location of natural resources such as minerals that support the energy transition[3], help develop a better understanding of our planet through deep geological time and assist with the prediction / susceptibility of areas to natural hazards such as earthquakes, landslides, and volcanic eruptions.

The use of AI is not new for the geosciences; the field has been at the forefront of using Big Data and AI methods. This includes the processing of massive seismic datasets, analysis of remote sensing data, and application of Natural Language Processing (NLP)[4] to millions of text articles - practices that have been undertaken in the geosciences for over a decade. Frontier AI such as Generative AI[5] presents further opportunities as well as amplifying existing risks and creating new ones. Such risks include loss of privacy and surveillance, bias and discrimination, concentration of power, fraud, loss of human autonomy and overreliance, and the spread of misinformation. Such risks present threats to science and human rights.

Ethical AI frameworks from bodies such as UNESCO[6] provide a foundation for addressing concerns including techno-solutionism[7]. These frameworks are, however, high level and lack the detail necessary for practical implementation in the geosciences context. This is evidenced by the misalignments, both past and present, between the way some geoscience AI initiatives are being designed, trained, deployed, and communicated, and the core ethical principles that should guide them. These have informed the prioritisation based on risk, for some of the recommendations in this report.

To address these misalignments, the Task Group on AI in Geosciences was established by the Commission on Geoethics of the International Union of Geological Sciences (IUGS) in November 2024. The mandate was to develop a set of practical and actionable recommendations for the ethical use of AI in the geosciences. By promoting responsible practices for 'social good', the Task Group seeks to ensure that AI applications in the international geoscience community

---

[1] https://www.unesco.org/en/artificial-intelligence
[2] https://sdgs.un.org/goals
[3] https://doi.org/10.1016/j.geogeo.2025.100361
[4] https://ieeexplore.ieee.org/document/9765270
[5] https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research
[6] https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
[7] https://library.oapen.org/bitstream/handle/20.500.12657/62281/1/9781000886078.pdf

realise their potential in ways that uphold universal human values, human rights, scientific integrity, and environmental sustainability. A glossary of terms (Appendix I), an overview of the philosophical framework guiding this work i.e., critical realism and virtue ethics (Appendix II) and methodology (Appendix III) provide details on the robust approach taken for compiling this report.

This report is intended to guide practicing geoscientists and those designing AI systems for geoscientists, needing clear, actionable guidance on responsible AI use. The one-page summary and illustrative geoscience case study examples are designed to help immediate practice. The report is also intended for policy bodies to highlight structural areas where concerns exist and the need for more considered approaches, standards and regulations aligned to UNESCO, the International Science Council (ISC) and jurisdictional AI laws. Finally, the report contributes to the academic discourse for AI ethics and theory development.

This document will be periodically updated to reflect the rapid technological changes in the AI landscape, including ongoing reflections on its ethical and social implications. The following section introduces the recommendations.

ETHICS RECOMMENDATIONS REGARDING THE USE OF AI

The following eight recommendations provide guidance on 'what good looks like' in the geosciences for geoscientists, policymakers, organisations/institutions and the general public. The recommendations (Table 1) form a basis for what should be considered, and which actions should be taken.

Table 1 - Recommendation areas for AI in the Geosciences

No	Recommendation
1	Use AI Responsibly
2	Promote Transparency and Explainability
3	Consider Bias and Fairness
4	Obtain Informed Consent, Protect Personal Data
5	Practice Participatory Design and Community Engagement
6	Advocate for Environmental Protection
7	Integrity in Science, Publishing and Education
8	Consider Geopolitics

A very simple one-page guide is presented for geoscientists, and those that build AI systems for geoscientists, that operationalises these recommendations for easy access and consumption.

Each recommendation is then discussed at length using virtue ethics (how to think about these issues rather than what to do) recognising the contextual nature of many of these issues. A criticism of virtue ethics can be how to use it in practice. To address this point, we have translated it into the following:

- behaviours (e.g., *Integrity – disclose AI use in your work*),
- reflective questions (*Am I communicating clearly to stakeholders when AI-generated outputs lack verifiable references?*) and
- scenario style "training" through the contextual geoscience illustrative examples (e.g., *A geoscience company uses satellite imagery and a proprietary AI/ML model to identify remote and rural areas for electrification…*).

A high-level roadmap is then presented for policy makers, to address further work identified in the report and to develop further resources, skills, and capabilities in this area.

One page quick-look AI Ethics guide for geoscience

Virtue Ethics Lens: These recommendations are not just rules, but reminders to act with responsibility, humility, justice, foresight, and integrity. Ask yourself: *What would a good and wise geoscientist do?*

1. Use AI Responsibly

- **Peaceful use of Geoscience AI should not be used to harm people or ecosystems.**
- **AI should be treated as a tool to support a geoscientists judgment, not replace it.**
- **Human oversight is mandatory where geoscience impacts people and ecosystems.**

2. Promote Transparency and Explainability

- **Research conducted with AI needs to be open, traceable and reproducible.**
- **Non-commercial models, training data, code & test sets should be open-source.**
- **Use explainable AI techniques to quantify uncertainty in black-box models and make limitations very clear.**
- **Use Retrieval Augmented Generation (RAG) to link Gen AI answers to sources.**

3. Consider Bias and Fairness

- **Use diverse, representative AI training datasets. Intervene where biases may impact the rights of marginalised / indigenous communities.**
- **No censorship in international facing geoscience AI outside the exceptions in the UN Declaration of Human Rights.**
- **Free web/cloud hosted AI tools aid fairness, but open-source tools best protect equity and data sovereignty.**

4. Obtain Informed Consent, Protect Personal Data

- **Consent to use personal data of any kind must be sought.**
- **Use opt-out mechanisms with providers when uploading sensitive data to hosted AI, or don't use them at all. Advocate for on-premise, / secure AI tools.**
- **Build AI systems with privacy-by-design, opt-in required for data use (e.g. model training).**

5. Practice Participatory Design and Community Engagement

- **Where AI outputs using earth science data (especially at scale using remote sensing data) impacts local indigenous communities – they must be meaningfully engaged "*nothing about us without us*".**
- **Adopt the CARE framework for indigenous data governance.**

6. Advocate for Environmental Protection

- **Geoscience institutions should support initiatives (e.g., UNEP) on AI's environmental costs and provider accountability.**
- **Advocate for sustainable practices (e.g. energy-efficient algorithms, renewable-powered data centres).**
- **Proactively look for opportunities where geoscience AI can aid env. protection.**

7. Integrity in Science/Education

- **LLMs can help refine your writing, but avoid 'cut and paste', use *your own voice* and *your own ideas*.**
- **Disclose and acknowledge all AI use.**
- **Gen AI can be "*more about persuasion than truth*". Be sceptical, check outputs and verify sources.**
- **LLMs can help as a coding assistant but not a replacement for learning how to code. Live explanations / oral defense can help reduce overreliance.**

8. Consider Geopolitics

- **International scientific / geoscientific institutions should remain neutral and avoid endorsing cloud/web-hosted geoscience Big Data & AI platforms that centralise data, as this risks eroding data sovereignty, trust, and reinforcing inequities ('algorithmic colonisation').**
- **Military ties in peaceful geoscience AI projects must be declared, ensuring informed consent for participation.**

1. Use AI Responsibly

Geoscience AI ethics should prioritize human rights[8], liberties and the environment, over efficiency. This is supported by UNESCO's AI Ethics Recommendations, "*Human rights and fundamental freedoms must be respected, protected and promoted throughout the life cycle of AI systems*".

Use cases for an AI system should be made clear, considering how they may be used at scale and in unexpected, or potentially harmful ways. Human oversight is necessary in AI-driven decisions, to validate outputs and ensure information aligns where appropriate with established scientific knowledge. **AI should be treated as a tool to support a geoscientist's judgement, rather than replace it**. Fully autonomous decision making in the geosciences is to be avoided where it impacts people, communities, or ecosystems. Geoscientists should be the guiding authority.[9]

An example where AI-driven decisions can impact communities unfairly if ethical human intervention is absent is given below[10] (see footnote).

Ethical issues of environmental sustainability versus socio-economic interests of poor mining families need to be balanced when implementing AI-based assessment of such activities. Notably, AI-based monitoring and reporting of artisanal and small-scale mining (ASM) do not necessarily lead to environmental improvements but may lead to the loss of livelihoods or to displacement of people. Governments, despite AI-based reporting, may continue to tolerate ASM due to its contribution to national GDP.

In the illustration above, remote sensing software or conventional statistical AI tends to require more human supervision and domain expertise. Frontier AI can automate pattern recognition and reporting across massive datasets, can more easily be repurposed, with little human input, which makes them more scalable and therefore potentially amplifies ethical risks if misused. Frontier AI outputs are increasingly being fed into live online dashboards and policy/decision making platforms. They are used by various actors, including government and conservation agencies, due to their scale and appeal to non-specialists.[11] They are now being used more extensively than advisory outputs from traditional statistical AI tools ever were. Heavy or blind reliance on AI systems, whether frontier AI or otherwise (e.g.

[8] https://www.un.org/en/about-us/universal-declaration-of-human-rights
[9] https://geoscientistscanada.ca/source/pubs/Insights%20into%20the%20Usage%20of%20AI%20Tools%20for%20Professional%20Geoscientists%20202509-%20Final.pdf
[10] https://ieeexplore.ieee.org/document/9954451
[11] https://www.maaprogram.org/amazon-mining-2024

due to automation bias), can undermine human accountability and make it easier to evade legal liability.

As AI systems are constrained by their training data, a plurality of AI systems should be guaranteed in the geosciences to avoid decision makers being unduly influenced by one set of training data. Increasing the responsible collection of data from parts of the world where there is a dearth of data (The Global Majority) whilst retaining privacy and data sovereignty needs careful ethical practice.

Recommendation 1: Use AI Responsibly

Additional reflective questions:

1. Am I embodying the virtue of responsibility by ensuring that I remain ethically and scientifically accountable for the decisions and outcomes produced through the use of AI?
2. Am I acting with integrity using geoscience AI for good not harm?
3. Am I demonstrating integrity and taking responsibility by not blindly using the recommendations from AI algorithms to make decisions that might negatively impact communities?

The following sections cover ethics recommendations throughout the AI lifecycle: Data collection and preparation, model development and training, model evaluation, adjustment, AI application development, deployment, and use.

2. Promote Transparency and Explainability

Transparency and explainability are concerns in the use of any technology. However, AI can produce systemic risks at scale, geoscientists may be less trusting of deep learning algorithms and models due to their opacity, whilst perhaps being too trusting of generative outputs from LLMs. In this section, and the following section on bias and fairness, we address these and other concerns.

Transparency

Open Science[12] in the context of AI is based on making resultant technologies freely accessible (Open Access[13]), transparent, reusable, explainable and reproducible. This is necessary to build trust in the outputs[14]. Documentation of the rationale for data used, design choices, parameters, failure modes, test sets, limitations and uncertainties are key for transparency and reproducibility. A risk-based approach could be taken prioritising those AI systems that directly impact people and ecosystems (by law in some jurisdictions)[15]. In this context, it is essential to make transparent the origins and gaps in the training data, its limitations, the associated uncertainties, data coverage and uncertainty across regions, and how these are quantified. Making data public, potentially utilising existing free earth science data repositories such as Pangaea[16], could be a low cost and fast way to release data.

In order for geoscience AI to be reproducible, the *actual* training data used (not a list of sources or metadata) should be released, with legitimate exceptions such as instances involving the right to privacy, safety, IP, indigenous sovereignty, and adhering to FAIR[17] principles (Findable, Accessible, Interoperable and Reusable).

Fully open-source LLMs exist outside the geosciences as an exemplar for open-source and Open science[18]. However, no examples exist in the geosciences. Openly releasing LLM model weights is beneficial. However, without the full open release of all underlying training data and training code, biases will remain opaque. See example illustration below:

A geoscience company uses satellite imagery and a proprietary AI/ML model to identify remote and rural areas for electrification. The company plans to license this system and related services to governments worldwide. However, the training data is not representative of all the regions where the system will be deployed. The AI model selects areas based on economic efficiency: prioritizing locations where electrification can be implemented quickly, easily, and at low cost to help governments meet targets. It does not consider actual need or urgency. In this context, it is essential to make transparent the origins and gaps in the training data, its limitations, the associated uncertainties across regions, and how these are quantified. The criteria for selecting certain areas, and excluding others, must also be clearly disclosed.

---

[12] https://www.unesco.org/en/articles/making-ai-more-open-could-accelerate-research-and-tech-transfer
[13] https://www.unesco.org/en/open-access
[14] https://www.nature.com/articles/s41561-024-01572-5
[15] https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence
[16] https://www.pangaea.de/
[17] https://www.nature.com/articles/sdata201618
[18] https://allenai.org/olmo

Funding bodies could stipulate data release requirements, and review assistance could be provided by standing geoscience ethical bodies (e.g. see high-level roadmap on page 28). Journal editors could also encourage certain AI training and test data is made available for reproducibility before manuscript publishing.

Where issues of IP come into play, training data, methods and code should be made available for regulatory scrutiny, which could involve standing geoscience ethical bodies providing domain expertise. This could ensure AI models are not compiled or used in violation of legal terms or ethical principles.

More ethical large scale AI text training data are now emerging[19] which do not infringe on an individual's IP and copyright, that could be adopted in the geosciences. Open-source software helps reduce barriers for under-represented communities, increases equity and democratises AI capabilities outside of technologically advanced enterprises and countries[20]. No code[21], or low code AI solutions can also help avoid discrimination based on background, age, gender, geography and ethnicity. If any components in an AI based application are not released and remain proprietary, the application cannot be reproduced. Geoscientists should display honesty and scientific integrity by being transparent and clear with their public communications on what is proprietary to avoid risks of "open washing" and "ethics washing" by releasing some data but retaining control.

LLM-driven applications such as chatbots should link to their sources for transparency, so any AI generated assertions can be checked at source by geoscientists[22]. This can mitigate the spread of misinformation and *hallucinations*[23]. An example is provided below to show issues can arise when this is not done.

A Chatbot was used by mineralogists researching minerals with diacritical marks. The chatbot suggested 'Eötvösite', a calcium magnesium iron phosphate, indicating the mineral was discovered in 1950 in Hungary, usually found in the oxidized zone of copper-uranium deposits, forming yellow-green crystals, also found in the Markey Mine in Arizona and the Rossing Mine in Namibia. However, after extensive checking by geoscientists, which was not easy as no online references were provided by the Chatbot, this was found to be a hallucination. The mineral does not exist. This illustrates how AI system design can hamper verification, and the importance of critical thinking.

[19] https://blog.eleuther.ai/common-pile/
[20] https://direct.mit.edu/itgg/article/1/3/119/9456/Open-Standards-Open-Source-and-Open-Innovation
[21] https://www.techtarget.com/searchsoftwarequality/definition/no-code
[22] https://www.journalofgeoethics.eu/index.php/jgsg/article/view/63
[23] https://arxiv.org/abs/2311.05232

To mitigate misinformation and errors in generative AI, geoscientific applications could adopt Retrieval-Augmented Generation (RAG) [24] or similar methods for reliable data sources that link AI outputs to verifiable sources. Without source traceability, geoscientists cannot reliably evaluate the provenance for explainability, uncertainty, or validity of AI-generated content. Explicit citation or linkage to underlying references should be the standard for any AI outputs used in research, education, or decision-making. These still need to be verified by geoscientists.

AI systems are increasingly being designed using a hybrid approach - using RAG but where RAG databases do not contain the relevant information, falling back to answers from within the black-box LLM. This must be made very clear to the user otherwise they may be misled and use some information which cannot be verified.

Explainability

All AI predictions carry a level of uncertainty and come with a set of limitations. It is important that researchers clearly communicate these especially when dealing with more opaque deep learning models. This is crucial when AI models are used in critical, high-stakes or time sensitive situations such as disaster preparedness[25] and may be required by law in some jurisdictions. Humanitarian organisations may rely on these models to plan for emergency responses where lives could be at risk. In such cases specifically, understanding and quantifying the uncertainty behind the predictions is essential and progress is being made in the geosciences on

Recommendation 2: Promote Transparency and Explainability

Additional reflective questions:

1. Do I ensure that the AI tools I use provide traceable references for their outputs, allowing me to verify scientific accuracy and provenance?

2. Am I relying on AI systems that obscure their information sources, or am I prioritizing tools that use RAG or similar methods to enable transparent review?

3. Am I communicating clearly to stakeholders when AI-generated outputs lack verifiable references, so that decisions are not made on the basis of unverifiable assertions?

[24] https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000877
[25] https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Explainable AI (XAI) for black box models [26-27] which can build trust and increase uptake. However, for deep learning and LLMs this is in its infancy, so we must recognise limitations. Systems are however, emerging to measure users' experiences interacting with explainable AI[28] that the geosciences may benefit from.

Scientific bodies across the world (such as the American Geophysical Union (AGU)) have developed Codes of Conduct for AI/ML researchers[29], including many of the aspects already covered. Where AI is considered a 'black box', this poses risk to both geoscientists and the general public basing decisions on outputs. Geoscientists should be able to validate AI systems at the model, parameter and algorithm level. This may create a stratification between those geoscientists who are AI savvy, and those that are not, and it is necessary to be mindful of the disproportionate influence this may have over decision making.

The aim should always be to generate explainable systems', covering such aspects as data provenance, feature importance and model cards.

3. Consider Bias and Fairness

Bias

Bias and Fairness in AI systems[30] is a multifaceted phenomenon. This ranges from biases in the training data collected and used to build AI models, to bias in algorithmic parameters and resulting applications.

There are existing standards that provide some guidance on Data Governance and AI design, such as ISO 8000-1:2022 "Data Quality"[31] and IEEE Standard 7003-2024 "Algorithmic Bias Considerations"[32]. In addition, ISO/IEC 42001:2023[33] is an international standard that specifies requirements for establishing, implementing, maintaining, and continually improving an AI Management System (AIMS).

Data-driven Language Models can help yield new scientific discoveries[34]. Conversely, AI primarily trained on scientific articles and popular publications, risks

---

[26] https://link.springer.com/article/10.1007/s10462-025-11165-2
[27] https://www.ecmwf.int/en/about/media-centre/focus/2025/unlocking-black-box-potential-explainable-ai-geoscience
[28] https://www.xeqscale.org/
[29] https://essopenarchive.org/users/536571/articles/635008-ethical-and-responsible-use-of-ai-ml-in-the-earth-space-and-environmental-sciences?commit=9a87c2b0c4ebfc3588ec1d515bc06fec3b1896a6
[30] https://www.mdpi.com/2413-4155/6/1/3
[31] https://www.iso.org/standard/81745.html
[32] https://ieeexplore.ieee.org/document/10851955
[33] https://www.iso.org/standard/81230.html
[34] https://www.nature.com/articles/s41586-019-1335-8?utm_campaign=related_content&utm_source=HEALTH&utm_medium=communities

reinforcing already established theories, further legitimizing them. This can lead to an "echo chamber" effect, where alternative or dissenting ideas are marginalized, not necessarily due to their lack of scientific merit, but because of their limited representation in the training data. In this sense, AI may contribute to preserving consensus than to fostering scientific innovation.

It is important to use representative, diverse and high-quality datasets that accurately represent the geographic, social, and economic contexts where the AI will be deployed. Structural biases may remain in human data, and it is important to raise awareness of in-built biases.

AI outputs should be monitored to detect and correct biases caused by omitted perspectives or censored information. As stated by the UNESCO AI Ethics recommendations, "*Member States should ensure that AI actors respect and promote freedom of expression*", fundamental human rights and scientific freedom should be respected by AI systems targeting the international geoscience community.

Guidelines for ethical content moderation that balance freedom of expression, regulatory compliance, indigenous rights, and user protection should be implemented. However, such moderation inevitably creates situational conflicts: what protects one group's rights may restrict another's, and local cultural or legal norms may diverge from global standards. This underscores the need for transparent, participatory processes (see recommendation 5) in developing and applying these frameworks.

Fairness

Given issues of geospatial data scarcity in some world regions[35], to ensure fairness it may be necessary for international agencies to provide financial support for responsible collection and curation of data from remote regions, especially where the private/public sector has little or no incentive to collect such data. Mitigation of unfair biases is a particular area of focus. An important mechanism for such mitigation and to ensure fairness are the CARE[36] Principles for Indigenous Data Governance (Collective benefit, Authority to control, Responsibility and Ethics).

---

[35] https://arxiv.org/pdf/2405.20868

[36] https://www.gida-global.org/care

The CARE principles guide the governance of indigenous data, ensuting it serves the collective benefit and rights of Indigenous Peoples. Existing power balances and historical influences should be factored into calls for adhering to FAIR principles so that these two sets of principles complement each other. An example of unfairness impacting indigenous or small farmers is given below.

Dilemma: AI-based assessment of climate-related agricultural risks can arguably lower overall risks for farmers and lower premiums for insurance. However, such AI systems may also reinforce existing inequities, as they may rely on incomplete or biased data that disadvantage small farmers with little digital or financial history. Such farmers may also be disadvantaged because of their traditional reliance on nature and natural farming methods, rather than on heavy chemical inputs. This can perpetuate cycles of poverty and exclusion from financial systems. Similarly, AI based assessment of satellite imagery that relies exclusively on past images to determine insurance premiums may also overlook present day shifts in agricultural practices that promise improved yields or better climate readiness.

Another example of bias and unfairness is shown below.

AI systems are suited to monitoring deforestation to protect vital rainforest ecosystems using ML algorithms and high-resolution satellite imagery. A hypothetical system identifies areas of rapid forest loss and flags them as illegal logging sites. The system was training on industrial-scale deforestation and lacked an understanding of traditional, community-driven, sustainable forestry practices. An Indigenous community has practised rotational agroforestry in a specific region for generations. This involves periodic clearing of small forest patches for cultivation followed by long regeneration periods. An inadequately trained and managed system cannot distinguish this sustainable practice from destructive clear-cutting and flags the community's activities as illegal. Government intervention restricts community access to their land and traditional practices. Livelihoods are disrupted, and traditional knowledge of sustainable land management is curbed and unable to be passed on to future generations. The unintended consequence is the loss of traditional knowledge and disrupted livelihoods, resulting in violations of intergenerational justice.

The unfair practices of some companies collecting AI training data have come under scrutiny for exploitative conditions.[37]

Free-to-use LLMs and AI applications have the potential to democratise AI, making advanced tools accessible “for the many, not the few” and helping to narrow the digital divide. By lowering cost barriers, geoscientists, small institutions, and communities can leverage AI for education, research, and innovation. However, such “democratization” can also lead to what some regulations call “systemic risks”.

Further, if the underlying training data or software code remain proprietary and are accessible only through cloud-hosted platforms, this can create new inequities, including the loss of data sovereignty and reduced local control over AI technologies.

Recommendation 3: Consider Bias and Fairness

Additional reflective questions:

1. Do I respect the human dignity of members of vulnerable communities?
2. Do I acknowledge their control over the knowledge products they create?
3. Am I ensuring that they receive material benefits when their data is used to train systems that generate decisions that may impact their future?
4. Do I practice compassion by considering how biased AI models in geoscience might disproportionately impact vulnerable populations, and am I taking steps to address or prevent that harm?
5. Am I implementing the virtue of justice by actively identifying, declaring and mitigating biases in the data, models and algorithms I use?
6. And am I committed to moral courage by questioning and challenging biased AI practices, even when it is professionally or institutionally inconvenient?

[37] https://www.theguardian.com/technology/2023/aug/02/ai-chatbot-training-human-toll-content-moderator-meta-openai

4. Obtain Informed Consent, Protect Personal Data

Personal data[38] include any information related to an identified or identifiable person. This also includes online identifiers such as the email address, Internet Protocol address, digital fingerprint, spatial proxies and geo-privacy in remotely sensed images and geospatial data. Many countries and legal jurisdictions (such as the European Union) have strict laws on how personal data should be handled and these laws must be followed. As stated by UNESCO recommendations on AI ethics, "*AI systems should not be used for social scoring or mass surveillance purposes*".

Rigorous anonymization and privacy-preserving techniques should be made when handling personal data of any kind (not just sensitive personal data). Accurate geolocated data such as maps interpreted by AI from remote sensing data on sensitive topics such as slums or refugee camps, may need to undergo privacy preserving techniques before being disclosed[39]. There are also increasing trends in AI research (machine vision in particular) to develop models that can be used for mass surveillance of humans, with striking patterns of obfuscation[40]. It is important for scientific honesty that motives and uses of AI are fully considered and disclosed.

Informed consent is a key aspect for personal data collection and use. Consent must be sought and explicitly granted before using data for AI training, requiring the consenting person to have the necessary knowledge on how their data will be used. For web/cloud hosted AI that centralises data harvesting, there should be clear and detailed Data Privacy and Cookie policies, including what personal information is being collected, whether it is depersonalised and if so how, where it is stored, for how long and for what purpose (including third-party access).

Many jurisdictions also have the 'right to be forgotten' in the Data Privacy laws for their citizens. However, this cannot be upheld once data has been used to train an AI model, which should be acknowledged.

AI tools widely used in the geosciences (e.g., language models, geospatial platforms, productivity software) often default to using user queries and uploads for further model training unless users actively opt out. This practice places an undue burden on geoscientists handling sensitive, proprietary, or community-owned data. Ethical AI use in the geosciences could adopt a privacy-by-design approach, requiring explicit opt-in consent before data can be harvested or reused for model training.

[38] https://commission.europa.eu/law/law-topic/data-protection/data-protection-explained_en
[39] https://arxiv.org/pdf/2405.20868
[40] https://www.nature.com/articles/s41586-025-08972-6

Terms of use for online AI should clearly explain the license of use and avoid vague or confusing language. They must specify what rights the AI provider and/or sponsor receives over the inputs (e.g., prompts and data uploads) and over the AI generated outputs. Users may upload data and documents on behalf of their institution, which could be confidential and/or of strategic importance particularly when related to natural resources. Whilst not 'personal information', the Terms should clearly inform users if third parties have access to the data, how it will be used and what national laws ultimately govern data access.

Recommendation 4: Obtain Informed Consent, Protect Personal Data

Additional reflective questions:

1. Am I ensuring that sensitive geoscientific data I enter into AI platforms is not inadvertently harvested for training by relying on default opt-out policies?

2. Do I require explicit, informed opt-in consent from collaborators or communities before uploading shared or proprietary data into AI systems?

3. Am I advocating for AI providers to implement privacy-by-design, so responsibility for safeguarding sensitive geoscience information is not shifted unfairly onto individual users?

5. Practice Participatory Design and Community Engagement

The involvement of stakeholders, including regulatory, end users of AI systems, and local communities impacted by predictions made by AI systems, is likely to lead to more ethical outcomes. This is especially true for cases when vulnerable or underrepresented communities that may be impacted are considered[41]. Funding agencies can require participatory research approaches where (potentially) affected communities contribute to data validation and AI model design. Involving those impacted may help ensure an accurate understanding of social contexts and their equitable incorporation into AI system design.

Public consultation and stakeholder inputs must be ensured before AI systems linked to land-use and provision of essential services, are deployed. Further, (potentially) affected communities should have clear legal recourse and

[41] https://www.sciencedirect.com/science/article/pii/S2666389921000155

compensation mechanisms if affected by decisions of an AI recommender system. An example is given below[42] (see footnote for quotations).

Scientists noted that Earth Observation (EO) research, when combined with AI-based predictions that replace traditional verification by ground truth, has a risk of leading to a kind of "colonialist science". As EO is, by definition, conducted remotely, "scientists who study places without ever going there, [may get] a false impression that they know the place and its circumstances. This can lead to misunderstanding or a complete lack of understanding of local contexts and may even lead to disempowerment or discrimination against local communities, their rights over ancestral lands or occupations, their cultural values, and even their means of livelihood."

Researchers should avoid making decisions using data without community representation, enforcing the guideline "*Nothing about us without us*": those affected must have the opportunity to participate.

Power dynamics dictate that technologically advanced and well-funded initiatives will shape AI deployment perhaps regardless of ethical guidelines. The challenge for fairness, is how early career geoscientists, developing nations and those less well-funded, *genuinely* participate and shape development. International geoscience institutions could play a key role here in facilitating this collaboration.

Recommendation 5: Practice Participatory Design and Community Engagement

Additional reflective questions:

1. Am I practicing the virtue of respect by genuinely valuing the knowledge, concerns, and lived experiences of communities affected by AI-driven geoscientific projects and allowing real participation by them in decisions?

2. Do I embody the virtue of epistemic justice by ensuring that all stakeholders, especially those historically marginalized, have an equitable voice in how AI tools are designed, deployed, and evaluated?

3. Am I cultivating the virtue of trustworthiness through transparent, ongoing relationships with communities, rather than engaging only when convenient or required?

[42] https://ieeexplore.ieee.org/document/10669817

6. Advocate for Environmental Protection

Data centres currently consume around 1.5% of the world's electricity supply, produce significant 'electronic waste', including hazardous substances like mercury and lead[43], and will account for one-tenth of global electricity demand growth by 2030[44]. The training and use of AI also uses significant amounts of water, and this environmental footprint depends on where this training takes place[45]. For geoscientists often described as Earth's custodians, AI use can pose dilemmas.

The use of AI can also have positive impacts on the environment. For example, global aviation contributes approximately 4% to global warming, with the clouds created by aircraft contrails accounting for 35% of all aviation warming impact. The use of AI on satellite imagery could be used to reduce this by 54%[46], so just this one-use case of AI could potentially reduce the world's global warming by 0.75%.

The geoscience community can lend their voice to the United Nations Environment Programme (UNEP), UNESCO and other bodies pushing for analysing the environmental costs of leveraging AI at scale. This includes the development of incentives and sustainable targets for AI providers, holding them accountable, to avoid adding to climate change. Measures may include energy-efficient AI algorithms, reusable or recyclable hardware, efficient cooling systems, use of renewable energy sources, and more energy-efficient computer chips.

AI systems can facilitate the optimal placement of data centres by identifying regions with stable renewable energy sources, which may help reduce environmental impact and improve energy efficiency. However, it has been reported some issues around environmentally friendly data centres are being downplayed[47].

Beyond this, AI also offers opportunities for geoscientists to preserve ecosystems and the climate. It can assist in the sustainable management of natural resources by helping analyse patterns, predicting risks, and guiding responsible decision making. Digital twinning, "... *advanced digital replications of complex and evolving systems using 'big' real-time data*"[48] has huge potential to improve risk identification in terms of infrastructure development and resource-consuming systems through smart urban planning for sustainable development[49].

---

[43] https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about
[44] https://www.iea.org/reports/energy-and-ai
[45] https://oecd.ai/en/wonk/how-much-water-does-ai-consume
[46] https://blog.google/technology/ai/ai-airlines-contrails-climate-change/
[47] https://www.mane.co.uk/resources/blog/the-green-illusion--the-rise-of-eco-friendly-data-centres-and-their-impact/
[48] https://www.sciencedirect.com/science/article/pii/S0160791X21002165
[49] https://www.nature.com/articles/s41559-020-01358-z

AI-based technologies applied in a number of research fields related to the environmental, sustainability and climate sciences are also gaining increased interest. These include AI applications in climate and Earth system modelling[50]; AI-augmented environmental monitoring[51]; autonomous underwater marine conservation interventions and data collection[52]; and AI-supported tracking of illegal wildlife trade[53].

These also come with potential harms, some of which have already been discussed in preceding sections. For example, algorithmic bias and allocation harm[54] through training data bias (limitations in terms of temporal coverage and geographical spread), transfer context bias (issues transferring AI systems from one ecological, climate or socio-ecological context to another – e.g., large industrial farm to small rural farming context) and interpretation bias (e.g., mismatching complex geospatial data and crop management action).

Recommendation 6: Advocate for Environmental Protection

Additional reflective questions:

1. Am I acting with environmental stewardship and temperance by carefully weighing the ecological costs of AI models, such as energy use and resource extraction, against their scientific benefits?

2. Do I embody the virtue of foresight by considering the long-term environmental consequences of deploying AI systems at scale in geoscience, rather than focusing solely on short term gains?

3. Am I showing integrity by advocating for more sustainable AI practices in my field, even if they require more effort or lead to slower computation?

7. Integrity in Science, Publishing and Education

Frontier AI poses new ethical challenges and opportunities for how we conduct geoscience, publish, write our research and reports, and how we educate current and future generation. AI literacy encompasses many of the areas already

[50] https://www.nature.com/articles/s41586-019-0912-1
[51] https://www.emerald.com/insight/content/doi/10.1108/978-1-83549-272-720241005/full/html?skipTracking=true
[52] https://inria.hal.science/hal-03686006/document
[53] https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.13104
[54] https://www.sciencedirect.com/science/article/abs/pii/S221204162030125X

discussed, moving beyond technical proficiency to incorporate critical thinking – informed, responsible, and adaptive. Geoscientists must be trained to maintain strict separation between training and representative test datasets to ensure scientific honesty, avoid bias and artificially inflated accuracy results. Efforts must also be made to apply robust validation techniques to ensure AI models perform reliably on unseen data and are tested against benchmarks. Without education in the foundations of data science, geoscientists may use AI techniques to generate biased or flawed results unintentionally, or some geoscientists may dismiss the use of data science and AI altogether[55].

Checking AI outputs rigorously

When using LLMs, AI-generated software code, images and textual assertions should be verified thoroughly, and textual sources checked diligently. AI-generated assertions can appear very convincing and plausible, especially when a reference link is provided. However, it is not uncommon for references to be fake or inaccurate for the AI-generated assertion. Outputs from LLMs can "*prioritize rhetoric over truth. They mix true, false, and ambiguous statements in ways that make it difficult to distinguish which is which. AI sounds convincing even when it's wrong. As such, current AI is more about persuasion than about truth. This is a problem because it means AI produces faulty and ignorant results. For now, we need to be highly sceptical of AI... The biggest risk is, as usual, ourselves*". [56] It is therefore vital that geoscience AI-generated outputs are verified by geoscientists, otherwise there is a risk of including flawed data and recommendations, or misinformation in research manuscripts and outputs. These may go unnoticed and undermine the scientific record, as well as "poisoning" future AI models trained on scientific literature.

Publishing

Further, AI tools cannot be listed as an author to a paper as a tool cannot take responsibility for the work[57]. As non-legal entities, AI tools cannot state the presence or absence of conflicts of interest nor manage copyright and license agreements. Where an AI tool is listed as a co-author (against the guidelines of most journals), all human authors must take responsibility for any resulting errors or misrepresentations.

Geoscientists who use AI tools to gather research data and/or help write the text of a paper or generate images, must disclose what tool(s) they used, how they were used and for what purpose. Many geoscience institutions (e.g., The Geological

---

[55] https://www.nrcs.usda.gov/sites/default/files/2023-03/Machine%20learning%20in%20Earth%20and%20environmental.pdf
[56] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5119382
[57] https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools

Society of London[58]) and Scientific Journals have published guidelines on AI tool use and are supportive of contributors using AI to help improve the readability of papers which can help individuals who are writing a paper in a language not native to them or have disabilities. This is as long as they adhere to their ethics, legal and best practice policies. The author is always responsible and liable for the work.

Geoscience is a highly observational and visual discipline reliant on imagery. For example, satellite imagery, outcrop photographs, 2D and 3D seismic and geophysical images, 2D and 3D cross sections and models, contour maps, time-depth borehole log electrical profiles, SEM images etc. Frontier AI techniques make it increasingly easier to manipulate imagery in very realistic ways. It may be deemed acceptable to use AI to remove a fence or road from an outcrop photograph using AI-based editing (although this should be declared). However, changing geological strata, mineral textures, or geological features in ways that modify original data (without declaring what has been done) could mislead and lead to serious ethical concerns, regardless of intent. Scientific integrity and trustworthiness in an era of frontier AI may require additional education on what is deemed ethical.

It is suggested that images should be labelled visibly, including on the image itself, where they are AI generated. Cut and paste of very large chunks of AI-generated text and left as-is, should be avoided in research papers – text should be *in your own voice* and convey *your own* original ideas. Otherwise, it's a form of plagiarism unless explicitly cited as AI Generated content. Geoscientific institutions are seeing misuse of AI in this context driven by time pressures and ease of use[59]. There are some tools[60] with reported high detection rates for AI-generated text, and AI generated text altered by humans, which may help support ethical codes.

Geoscience journal editors receiving papers with an AI component, should ensure those with the relevant expertise are invited to peer review submissions. This may help ensure training data is released in full (or else documented as proprietary), training and test methods documented, and models and code released publicly.

Education

As stated by UNESCO in a report published in September 2025 'AI and the future of education: disruptions, dilemmas, and directions'[61] the "*growing presence of generative AI marks another major historical rupture in education*". This brings ethical concerns, "*risk cognitive offloading, decline in critical thinking and reinforce divides*

---

[58] https://www.lyellcollection.org/publishing-hub/ai-in-publishing
[59] https://www.grss-ieee.org/resources/news/grss-addresses-the-benefits-and-pitfalls-of-ai-in-preparing-research-papers-for-publication/
[60] https://www.pangram.com/
[61] https://unesdoc.unesco.org/ark:/48223/pf0000395236

of access, gender and language. As they generate data at unprecedented scale, they also raise unanswered questions of privacy, ethics, sovereignty and trust".

The use of frontier AI such as LLMs in geoscience education and outreach is an emerging area in science[62]. LLMs can also help software coding, but to avoid 'shallow learning', assessments may need to be modified to avoid overreliance, perhaps through live explanations / oral defense of software code or other techniques. One of the authors of this report has just started liaising with a university in the Global South to incorporate LLMs into their geoscience curriculum and is an area of further work.

In geoscience specifically, the proliferation of highly realistic completely AI-generated images presents risks: members of the public, and even trained scientists, may mistake fabricated geological scenes for authentic field photographs[63]. Similarly, new AI-driven platforms[64] that claim to "score" geoscience papers or reports raise questions about fairness, bias, and accountability.

Given these dynamics, this is an area where an AI ethics standing body in geoscience (see roadmap recommendations) could play a valuable role. Such a

Recommendation 7: Integrity in Science, Publishing and Education

Additional reflective questions:

1. Am I demonstrating intellectual honesty and integrity by clearly disclosing how, where, and why AI tools were used in my research and manuscripts, and have I checked outputs and references? Do I cultivate the virtue of humility and integrity ensuring my manuscript and research outputs (inc. peer review) are my own ideas and in my own voice?

2. Do I cultivate the virtue of humility and integrity by acknowledging as a geoscientist the limits of my own understanding of statistics and data science, taking appropriate action so as not to inadvertently create biases within my research, models and algorithms?

3. Do I demonstrate integrity, humility and scientific honesty by avoiding exaggerating or inflating the results or benefits of AI from my work or research, or the work of others or technologies I am associated with?

---

[62] https://www.sciencedirect.com/science/article/pii/S1041608024001948
[63] https://x.com/GeologyPage/status/1924203073506832541
[64] https://geologyoracle.com/

group could help set standards for responsible use, promote transparency in educational applications and tools to help verification of AI generated outputs, and guide how geoscience educators and communicators harness these technologies without compromising trust or scientific integrity. By providing a focus on resources and awareness, smaller institutions and organisations, which may not have the economic resources, could utilise these resources for their members and staff.

8. Consider Geopolitics

Considering the impact of power dynamics on AI ethics, Geopolitics is a key element, especially for LLM and LLM-based applications. As highlighted by the International Science Council (ISC) [65], areas which should be given particular consideration include:

- *Is a desire for technological sovereignty driving behaviours?*
- *Digital Colonialism: Could state or non-state actors harness systems and data to understand and control other countries' ecosystems?*
- *Digital Divide: Are existing digital inequalities exacerbated, or new ones created?*
- *Military: Is there a possibility for both military application as well as civilian use?*

There are useful examples of free frontier AI tools that allow geoscientists worldwide to access and interpret data. This aligns with UNESCO's recommendations on AI ethics, "*justice, trust and fairness must be upheld so that no country and no one should be left behind*". It is important to foster democratisation of AI capabilities, considering that AI can further concentrate power (AI and data) limiting it to a small number of actors, further exacerbating digital and learning equalities.

However, when mediated through web/cloud-hosted AI, it is often conditional on granting rights over data use and storage. This raises significant concerns about power imbalances, transparency, and consent. Stakeholders in AI systems should be aware of the impact of the geopolitical powerplay on the Global South. As highlighted by the ISC, AI systems that enable *algorithmic colonialization* risk reinforcing dependency and unequal power dynamics and should be avoided.

Using web/cloud AI requires users to upload their own data, which centralises mass data collection. Whoever hosts the AI (independent of what AI/LLM models are used) – controls the data. Even the original platform provider may lose control of the intended data use. This is because platform terms of use can be superseded by the legal framework of the hosting country or the laws that govern the legal entity operating the platform. A UNESCO report on AI education[66] notes that AI is becoming, "*increasingly an instrument of statecraft for countries like the United States of America and the People's Republic of China who invest in AI frontier*

[65] https://council.science/publications/framework-digital-technologies/
[66] https://unesdoc.unesco.org/ark:/48223/pf0000395236

research, infrastructure and cybersecurity with the same foresight and strategic intent reserved for military and economic power".

Geoscience intersects directly with national security touching on critical minerals, oil and gas, water, and other natural resources. For these reasons it is recommended that international scientific and geoscientific institutions:

- **Do not lend their name, support or affiliation** - either directly or indirectly, to initiatives or proprietors that use web/cloud hosted AI platforms (or has an MoU to promote such tools) which centralise information harvesting of geoscience data from individuals, institutions and countries. Doing so risks legitimising the erosion of data sovereignty and politicising scientific institutions, thereby fracturing trust within the international community.
- **Remain neutral** – in this highly competitive and geopolitically sensitive domain. Rather than endorsing specific AI technologies, AI providers or Big Data and AI initiatives, scientific and geoscientific institutions should prioritise skills development, AI literacy and support fully open-source initiatives.

Following this approach could help enable all geoscientists and countries to develop their own capabilities, autonomy and control of their data. Individual geoscientists and institutions are free to use web/cloud hosted AI from any proprietor, open-source alternatives, or develop their own. Neutrality of international institutions is, however, essential to safeguard trust and equity.

Power asymmetries may also exist between regulated industries in the geoscience sector (such as oil and gas, mining, renewables etc.) and government regulators. AI

Recommendation 8: Consider Geopolitics

Additional reflective questions:

1. Am I acting with foresight, helping build sustainable local capabilities in AI, recognising where new detrimental dependencies are being created?

2. Am I acting with honesty and integrity communicating all the underlying geopolitical motives for geoscience AI based research projects?

3. While participating in civilian geoscience spaces, am I respecting the moral autonomy of the geoscience community by fully disclosing my ties to military research, knowing that others may not want to be associated with such aims?

can help enhance regulatory quality, and reduce compliance loads, but use can also cross over into weakening oversight. The ethical use of AI must explicitly address its role in shaping regulatory processes in these sensitive sectors. AI can exacerbate power asymmetries, where regulated industries often hold more technical expertise and information than regulators. By exploiting large datasets of past regulatory decisions and profiling regulators' preferences, companies could deploy AI to anticipate, influence, or even manipulate regulatory outcomes. This raises the risk of regulatory capture, where AI becomes a tool to distort decisions in favour of powerful stakeholders, undermining fairness, and public interest.

HIGH-LEVEL ROADMAP

The roadmap to address and implement these recommendations is twofold. Firstly, to develop practical ethical impact and risk assessments, and secondly to establish a standing body in the geosciences to take some of these recommendations further towards implementation.

(i) Governance through Ethical Impact and Risk Assessments

Ethical impact and risk assessments should be undertaken for any AI study or deployment. It is suggested that this be done through a thoughtful critical realist and virtue ethics approach, rather than a tick-box exercise.

They should also be used to inform any project kick-off, go-live, or stop recommendations. Where possible, these ethical risk and impact assessments and their considered justifications should be made public. The above eight recommendations can support such ethical risk and impact assessments.

1. Use AI Responsibly
2. Promote Transparency and Explainability
3. Consider Bias and Fairness
4. Obtain Consent and Protect Personal Data
5. Practice Participatory Design and Community Engagement
6. Advocate for Environmental Protection
7. Integrity in Science, Publishing and Education
8. Consider Geopolitics

There are some international templates for ethical risk assessments that can be used including those from UNESCO[67] which outline additional questions that should be considered. The MIT AI Risk Repository also provides an overview of AI threats and a frame of reference for practitioners[68]. See the next section for proposed recommendations to create geoscience specific template examples.

The pros and cons as well as competing explanations must be explicitly weighed for every aspect of an AI system. Harms are harms, intentional or not, and unintended outcomes can be worse than the initial issues being addressed.

Depending on the nature of the AI system, emphasis can be given appropriately. For example, where energy-intensive computing power is necessary (e.g., climate

[67] https://unesdoc.unesco.org/ark:/48223/pf0000386276
[68] https://airisk.mit.edu/

simulations, large-language model training) this will likely require more attention on the environmental section than small scale AI studies.

Particular attention could be paid to peer reviewed scientific research on AI ethics for EO [69,70,71] and LLMs in the geosciences[72] which highlight several AI-ethics issues that have occurred, are occurring, or have a particular affordance to occur in the future.

These ethical risk and impact assessments could help support ethical decision making in a considered, and equitable manner supported by education and expert independent advice. The next recommendation lays out how these ethical risk and impact assessment templates, education and support may be developed and made available to the geoscience community, along with other suggested recommendations to support ongoing developments.

(ii) Establish a standing body for AI ethics in the geosciences

AI is an incredibly fast-moving area of study and development where ethical issues and dilemmas are surfacing very rapidly, almost in real time. There is a desire from private technology vendors and government-funded initiatives to deploy AI systems as quickly as possible to gain first mover advantage, which outpaces regulatory development.

Issues around AI Ethics are only going to expand. The geosciences have unique characteristics that justify a dedicated approach rather than relying on generic AI ethics frameworks. The international nature of Earth observation, the intersection with national security and resource sovereignty, and the direct connection to climate and environmental policy create a distinct ethical landscape. There have been, and remain, repeated AI ethics failures in the geosciences.

It is proposed that a pragmatic standing body of volunteers be created within the IUGS Commission on Geoethics to further develop guidelines and resources that support the international geoscience community in implementing AI ethics. This should be transdisciplinary, including legal experts and social scientists as well as geoscientists and technical experts. Through the proposed standing body, activities could include (but not be limited to):

---

[69] https://ieeexplore.ieee.org/document/9954451
[70] https://arxiv.org/pdf/2405.20868
[71] https://ieeexplore.ieee.org/document/10897919
[72] https://www.journalofgeoethics.eu/index.php/jgsg/article/view/63

- Develop culturally inclusive educational tools and resources to raise the level of digital literacy in AI in an equitable way. If geoscientists at an early stage of their career do not get the opportunity to develop and use AI, they may be left out of labour markets.

- Promote human-AI collaboration in geoscience field activities and education, where AI augments but does not replace critical thinking, human-driven inquiry or human oversight and determination.

- Develop specific guidelines and codes of conduct around aspects of AI in the geosciences such as a template for ethical impact and risk assessments.

- Create a register for the environmental impact of AI models in the geosciences.

- Provide independent support and input into ethical risk and impact assessments for major initiatives in the geoscience community.

- Support geoscience editors and journals, where they feel they do not have the necessary networks for peer review of geoscience related AI papers.

- Provide ad hoc reports to geoscientific institutions/organisations on AI ethics topics when advice is sought.

- Collect benchmark test sets for geoscience AI to promote wider awareness and use, to help compare AI models with one another and help assess generalisability.

CONCLUSION

Ethics is not just about rules or consequences; it is situational, emotional-empathetic and relational. It is about moral character. Virtue ethics is a habitual disposition to act rightly – what a good and wise person would do.

The recommendations are intended to support the international geoscience community on 'what good looks like' for ethical AI. Some recommendations call for further work, while others are ready to be implemented immediately. These are grounded in the challenges faced by technology use and technology deployments in the community. These recommendations provide supporting detail to the AI Ethics guidelines from UNESCO.

The opportunity for AI in the geosciences is potentially transformational for both productivity and scientific discovery. There are also challenges and risks that could do significant harm to the international geoscience community, the general public, and our Earth and its resources, unless rigorous ethical guardrails are established.

ACKNOWLEDGEMENTS

Reviewers

Thank you to the reviewers, for providing excellent constructive feedback that we believe has significantly enhanced the quality and usefulness of the report.

Use of AI

Generative AI tools (LLMs) were used (OpenAI GPT-4, Anthropic Claude 3.5) to help prepare this manuscript. This included searching for relevant information, checking the report for inconsistencies, readability, and stimulating ideas (after initial brainstorming with human input only), and examples through 'what if' situational analysis which were subsequently modified.

APPENDIX I – GLOSSARY

Critical realism: Is a philosophical framework in which reality is taken to exist independently of human thought and perception. This "real" world is understood as not just a collection of observable phenomena but as having underlying structures and causal mechanisms that underpin the "observable" world. Critical realism highlights the interplay between social structures and individual agency; assuming that, as humans, we are both shaped by, and can act upon, our environment.

Fallibilistic epistemology: In critical realism, fallibilistic epistemology refers to the belief that all knowledge is uncertain, incomplete, and tentative, regardless of how well-supported it appears to be. This does not mean there is no justification for knowledge, but simply that there are constraints for making knowledge claims, as 'truth' is understood as 'warranted assertability'.

Judgemental rationality: Is the critical realist principle that we can rationally evaluate and compare competing knowledge claims, even though all knowledge is theory-laden and fallible. It states that we are not trapped in relativism, there are universal principles in AI ethics, and that some theories can be judged better or worse (more explanatory, coherent, empirically adequate) than others.

Warranted assertability: Refers to whether a particular claim or theory from competent scientific inquiry, can be justifiably asserted at a given point in time, based on current evidence, reasoning, and explanatory power.

Virtue ethics: Is an approach to morality focused on developing good character. Instead of just following rules or aiming for the best consequences, it asks what a good person would do, guided by moral virtues like honesty, courage and empathy, and intellectual virtues such as practical wisdom and intuition. In virtue ethics, the central question is "*What kind of person should I be?*".

Artificial Intelligence (AI): This technology allows machines to mimic human intelligence often using data whose scale exceeds what humans can analyse. Machine Learning (ML) is a type of AI method that develops outcomes through models that have been 'trained' from data patterns without human control or direction. 'Learning' refers to an algorithm's ability to update weights in a statistical calculation to improve the classifications and predictions it generates. AI systems directly impact the world, as through their models ML algorithms may determine who needs to be evacuated within an area from a predicted imminent landslide, credit risks, who may get a job, who might be stopped by the police, who may receive a particular type of medical treatment and what land may be developed etc. ML algorithms turn humans and communities into data subjects - into ranked and rated objects.

Models can be trained on labelled data (supervised machine learning) or without labels (unsupervised machine learning) and can be multi-modal such as text, structured data, images, video and audio. Deep Learning (DL) can be considered a subset of machine learning which uses many layers in an artificial neural network, for example Transformers, Large Language Models (LLMs) and Vision Language Models (VLMs). AI can also include techniques such as rule based systems and ontologies/taxonomies, as well as data driven ML techniques.

APPENDIX II - PHILOSOPHICAL FRAMEWORK

The nature of ethical decisions and human-technology interaction is contextual and therefore fallible in the sense of being updatable should context change. Different concerns arise in different sectors, while regions and individual communities are differently vulnerable to potential harm from AI technology. Perhaps most impactful is the rapid evolution of AI technology and its implications for human agency. Ethical decisions can be complicated; however critical realism offers a philosophical framework to help.

Critical realism[73], with its interpretive fallibilistic epistemology, creates the possibility for providing justifiable recommendations to address ethical questions. This philosophical approach provides the backdrop for a focus on virtue ethics[74] that emphasises both moral and intellectual virtues in deliberations needed for ethical decision making. This approach ensures rational justification while decisions remain context dependent.

The contextuality at issue in 'ethics of AI' discourse does not mean that we are caught in ethical relativism (although it does suit certain power narratives to imply that this is the case[75-76]). Instead, it highlights how, now more than ever, we need reasoned debate about what is right. We also need to consider what is the right way to prevent or respond to harm in every context. Finally, we need to look ahead and consider how contexts and perceptions might change when we reflect on appropriate ethical decisions in a given situation.

To address the challenge of justification, we need a different approach. A useful approach is to interpret decisions through the lens of warranted assertability as understood in critical realism. This means acknowledging that knowledge is not absolute. At the same time, we must ensure that our justifications are as objective as possible (e.g., experiments are public and repeatable[77]). A virtue ethics approach enriches this effort by emphasizing both intellectual and moral virtues. Intellectual virtues such as curiosity and intellectual honesty help guide open and thoughtful scientific inquiry. Moral virtues such as transparency and accountability can ensure ethical responsibility. Together, these virtues strengthen the justification given for ethical decisions in fluid and complex contexts.

To build just and responsible AI ecosystems where transparent and rational justifications can be given for the decisions reached, we need people with a

---

[73] See, Collier (1994).
[74] For details, see Aristotle's Nichomachean Ethics: https://historyofeconomicthought.mcmaster.ca/aristotle/Ethics.pdf
[75] https://blog.citp.princeton.edu/2019/03/25/ai-ethics-seven-traps/
[76] https://firstmonday.org/ojs/index.php/fm/article/view/13636/11606
[77] See, Peirce (1955).

particular kind of reasoning skill. This skill involves analysing ethical dilemmas based on two things: (1) the situation a person or a group/committee is in (relating to the need to act with moral virtues), and (2) the knowledge available to them (relating to the need to act with intellectual virtues). With these tools, they can address dilemmas and articulate their reasons for making certain choices at certain times, rationally and responsibly.

It also has to be kept in mind, as alluded to above, that whether a statement becomes or remains, a belief, or a viable decision, depends on more than the context in which it was made. It also depends on how it is judged in future contexts, as situations evolve. Changing contexts are all the more reason to focus on rational and reasoned justifications for decisions, rather than relying solely on universal rules.

APPENDIX III - METHODOLOGY

The scope of this study is AI ethics in the geosciences which includes both the geological sciences and the Earth Observation (EO) discipline. AI ethics overlaps with existing ethical concerns around technology and science, however, AI can significantly amplify these effects (such as data privacy and surveillance, authorship integrity), as well as generate new ones (bias in training data, explainability of complex AI models and autonomous decision making).

Frontier AI (e.g., trained on un-curated corpora, black-box large artificial neural network models, persuasive) can differ from traditional statistical based AI (e.g., regression, decisions trees, labelled small datasets, overfitting, inappropriate generalisation) – both are in scope of this study. General ethical questions about existing national or international laws pertaining to AI were placed outside the scope of this Task Group.

A literature review (including company reports, websites, social media, news and academic published papers) was undertaken as AI relates to the geosciences. Numerous search engines and LLMs from different countries were used to gather references, so not biased by geography, but biased by a publication bias. Searches were informed by using internationally accepted ethical AI principles (UNESCO[78]), AI analytical frameworks (International Science Council - ISC[79]) and emerging AI Laws and Regulations, to the geosciences.

These data were then triangulated with a longitudinal analysis of the empirical behaviour of AI technology systems, models and use in the geosciences, to ascertain how principles were often put into practice – including several detailed examples of ethical failures over time.

This was viewed through a lens of critical realism, which is particularly helpful as it forces us to look at hidden structures and power dynamics behind AI use in geoscience. It forces us to ask questions like "*Who gains power from this AI deployment?*" and "*Who loses agency?*". For example:

Misclassification or inappropriate classification of a geographical area is undesirable. For instance, arable land being classed as barren to suit the interests of mineral explorers who want access or suiting a government to extend infrastructure projects. AI should also not be used to obfuscate motives or remove individual accountability of certain parties for decisions made.

---

[78] https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence
[79] https://council.science/publications/framework-digital-technologies/

Critical realism places questions of power, inclusion, and exclusion central to ethical evaluation avoiding narrow, reductionist evaluations.

These current issues were supplemented with counterfactual "What if?" thinking about future ethical AI issues, accepting it is impossible to predict trends that might suddenly emerge, acknowledging some are more certain e.g. more satellites being launched collecting more high-resolution data. This was achieved by extrapolating current trends, including data collection, AI technology and Big Data to suggest future scenarios and areas of concern. An illustrative example is shown below:

The increase in numbers of satellites, Uncrewed Aerial Vehicles (UAV) and remote sensing data resolution, with more advanced AI detection, and the tendencies for making more data Open Access poses a range of ethical challenges in the future. These include violation of geo-privacy, and potential surveillance, impacting individuals and vulnerable communities, as well as global peace and security.

Areas of misalignment between internationally agreed ethical principles and what is actually happening, as well as future areas of concern, were clustered to form eight core areas where recommendations are needed. The authors represented 'Global North' and 'Global South' perspectives for a balanced and inclusive approach. Virtue ethics was used to inform the recommendations - an approach to morality focused on developing good character. Instead of just following rules or aiming for the best consequences, it asks what a good person would do, guided by moral virtues like honesty, courage and empathy, and intellectual virtues such as practical wisdom and intuition. This conceptual framework is shown in Fig. 1.

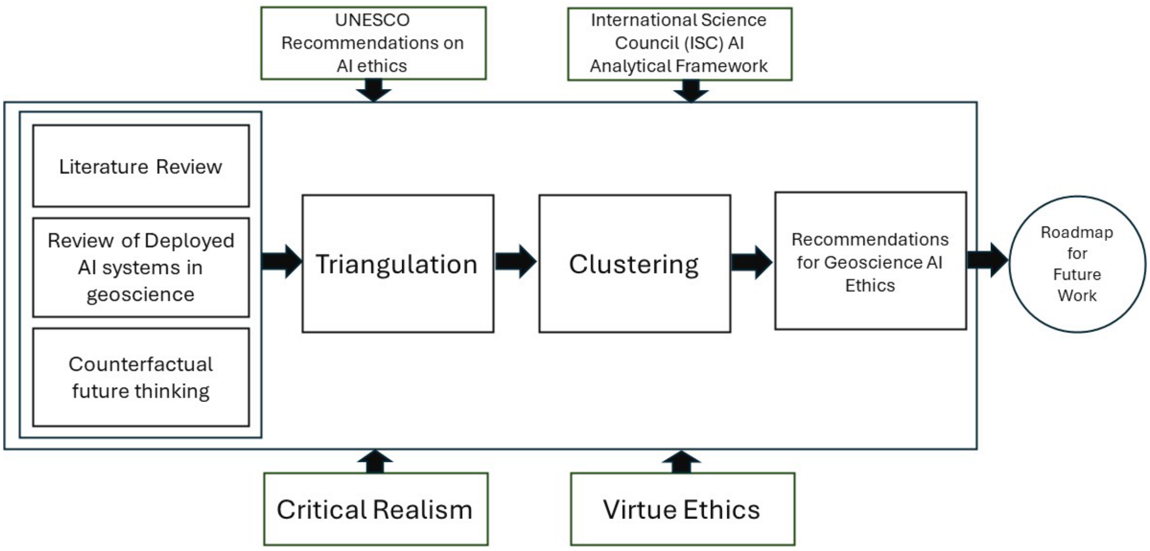

Figure 1. Conceptual Framework used by the Task Group on AI in Geosciences