2026 ITRC Project Team Proposal

Environmental Topic

Artificial Intelligence (AI) and Machine Learning (ML) in the Environmental Field

Proposed Project Title

Leveraging Artificial Intelligence and Machine Learning to Efficiently and Responsibly Manage and Analyze Environmental Data

Project Deliverables

The new team would begin in January 2026. Given the rapidly evolving nature of AI and ML technology and its broad-reaching impacts across industries, we propose a shorter term (12 to 18 months) team with a smaller set of project deliverables, mostly involving data gathering from ITRC members and partners and a 'state of the science' report. The findings of this report may help inform one or more proposal(s) for future teams on specific topics relevant to regulatory concerns (both state and federal) that are not already addressed by existing, publicly available AI/ML resources.

This project will deliver valuable information to regulators and the public, facilitate information gathering and sharing beyond the initial project scope, and may inform future projects related to AI/ML in the environmental field. Proposed ITRC work products include:

- ITRC Membership Survey: Soon after the team kicks off, a survey will be sent to state and federal agencies, Industry Affiliates Program (IAP) members, and other ITRC partners to assess their needs and understanding related to AI/ML, along with gaining a better understanding of how these technologies align with member information technology (IT) policies.
- One Focus Sheet: The initial deliverable from this team will follow a similar format to the paper developed by the 6PPD[1] (an emerging contaminant) team and will provide a high-level overview of available information on AI and ML, definitions of key terms related to the topic, and examples of applications in environmental data management and analysis. We expect participation with members of the former ITRC Environmental Data Management (EDM) team and plan to build upon topics addressed there, including data governance and quality. This focus sheet will be instrumental in helping regulators get up to speed on these technologies and gauge relevance to their regulatory work. While the use of AI and ML in the environmental field is simultaneously novel and accelerating rapidly, the focus sheet will provide a common starting point for information on the fundamentals and current status of AI and ML.
- Case Studies: Examples of effective AI and ML implementations in state, federal, and private sector environmental projects.
- Internet based training – Either a set of produced instructional videos, recordings of webinars, or videotaped presentations at a conference.

[1] https://6ppd.itrcweb.org/wp-content/uploads/2023/09/6PPD-Focus-Sheet-Web-Layout-9.pdf

Problem Statement

Federal and state environmental agencies and private sector organizations face significant challenges in managing and analyzing the growing volume of environmental data (e.g., multimedia sampling results, human health and ecological hazards, continuous monitoring data) without corresponding increases in staffing. Traditional methods of data management and analysis often struggle to handle the complexity and scale of modern datasets. Artificial Intelligence (AI), which simulates human intelligence, and Machine Learning (ML), a subset of AI that learns from data, offer powerful tools to enhance data management and analysis and facilitate decision-making, accelerating cleanups and approvals by saving time and decreasing costs. Going forward, this proposal will use the term AI to refer to both AI and ML technologies.

However, there is a lack of clear guidance and best practices for implementing these technologies to manage and evaluate environmental data. Further, regulatory agencies have a "duty to protect" the public from environmental hazards that affect humans and ecosystems. This duty adds a layer of complexity in incorporating AI as there may be real and perceived risks to relying on the results of analyses and decisions made by machines. The use of AI in environmental decision-making represents an example of high-impact AI, as it directly influences human health and safety.

Given the rapidly evolving nature of AI technologies, there is a strong need for material that explains the fundamentals of what these tools are and how they operate in a way that is approachable to a diverse audience. Given the rapid evolution of AI technologies, there's a risk that prepared documents may quickly become outdated. Nevertheless, ITRC members and partners currently need assistance in identifying effective ways to start integrating AI, as these technologies have become too widespread to overlook.

Background: AI is a rapidly evolving field with significant potential to transform environmental data management, analysis, and decision-making. These technologies can automate data processing, identify patterns and anomalies, and provide predictive insights, which might be missed by traditional methods. Despite its potential to accelerate data management and evaluation, many federal and state agencies and organizations struggle to understand and apply AI effectively.

Public and private sector environmental professionals are currently struggling to identify which direction to go with implementing AI solutions to address the challenge of managing and analyzing constantly increasing volumes of data with limited, if any, increases to staffing. Additionally, regulators and other environmental practitioners need to balance the need for increasing efficiency with high pressure for quality control given the potentially serious implications of even small errors (e.g., using an incorrect regulatory standard) for human and/or ecological health. Although automated outputs generally produce fewer errors than manually prepared deliverables, deployment of AI technologies requires human verification to ensure the information inputs are accurate and the proper methods are being applied, and to address quality control and ethics concerns (this is known as "human-in-the-loop" AI). Given these competing objectives, the effort necessary to find a viable path forward can pose a barrier for resource-constrained agencies.

Technical and Regulatory Barriers: This proposed team aims to address the following technical and regulatory barriers currently faced by ITRC members:

- High costs and slow turnaround times associated with management, analysis, and quality control of environmental data via traditional methods. Staffing challenges and budget limitations across

the environmental field are driving a need for increased productivity to keep up with expectations from regulators, the regulated community, and the public. At its best, AI offers a force-multiplier tool to increase productivity and efficiency while maintaining protectiveness of human health and the environment;
- Lack of understanding of the current state of AI within the environmental sector;
- Concerns about data privacy and security;
- Regulatory uncertainty and the need for clear guidelines on data governance, quality control, and transparency when using AI applications; and
- Limited access to AI tools and expertise within state agencies and smaller organizations.

Proposed Solutions:
To assist state, federal, and private sector organizations in overcoming the technical and regulatory barriers to modernizing their data infrastructure and leveraging AI applications, we propose the following solutions:

- Prepare a survey to assess ITRC members' needs and understanding related to AI. Possible survey question topics include:
    - Current roadblocks faced by members in implementing AI technologies;
    - Status of an AI Acceptable Use Policy or AI Guidelines;
    - A numeric self-rating of AI Maturity on a scale from 1: "No AI use" to 5: "Production-level AI implemented in multiple programs";
    - Alignment of AI technologies with member information technology policies and data infrastructure;
    - Languages (Python, R, etc.), libraries (scikit-learn, TensorFlow, PyTorch), and data platforms (SQL, GIS systems) currently in use;
    - Consideration of ethics when making decisions based on potentially biased training sets; and
    - Consideration of the carbon and water footprints of AI technologies.
- Develop one high-level focus sheet, similar to that prepared by the 6PPD team, to provide foundational knowledge on AI. The focus sheet will include the following:
    - An emphasis on the importance of drafting an AI Acceptable Use Policy or AI Guidelines for responsible development and use of AI with a recognition that an organization's AI guidance and strategy will need to continually evolve as organizational needs and goals, technology, and the state of the practice evolve;
    - A one-page flowchart depicting the general lifecycle of an AI model (data ingestion, feature engineering, model training, validation, deployment, and monitoring) to provide a clear visual anchor for those new to the topic;
    - A summary of available information on AI to help those within the environmental industry achieve a shared, baseline understanding of this technology;
    - A list of existing resources on AI, including practical guidance for modernizing digital infrastructure to seamlessly integrate AI tools with contemporary environmental data management schemas for those interested in taking a deeper dive;
    - A preliminary list of publicly available, open-source AI frameworks that may be built upon, which may be used to lower the costs of AI integration within state, federal, and private environmental organizations. To address cybersecurity concerns related to larger models using data for training, this list will also identify if there are smaller publicly available, open-source models that can be trained and run locally on more curated data sets; and
    - Preliminary guidelines on best practices such as the need to document and audit work prepared using AI tools to balance the quest for efficiency with quality control and risk management. The prior ITRC EDM Best Practices Team developed fact sheets related to environmental data quality, including verification and validation, which includes application of data qualifiers to convey information about quality issues to end users, and

data usability, which is a further level of review that evaluates the quality of the data for its intended use(s). Given the rapid evolution of AI, it is important that methods for managing and analyzing environmental data using AI are documented and well understood so that similar applications of data qualifications and usability assessments can be performed, in turn building end user trust in both the new methods and the usability of the data. It is anticipated that some of the best practices described in those existing fact sheets related to data quality and public confidence are applicable to AI outputs and may be built on by this team. To further enhance the reliability of AI-generated outputs, integrating a "human-in-the-loop" approach can ensure that critical decisions are informed by human oversight and judgment, fostering greater confidence in data quality and usability.

- Compile case studies to demonstrate applications of these tools to increase efficiency and reduce costs when managing and evaluating environmental data sets, whether effective or not. It is anticipated that these case studies and "lessons learned" will enhance confidence in these technologies and raise awareness of innovative applications to meet ITRC members' needs. Each case study will include a review of why the project was or was not effective (e.g., input data pre-processing, human intervention in the AI routine, simplicity of the data itself). Example use-cases that may be included in case studies include:
    - Identifying trends and patterns in remediation site data (e.g., source differentiation and forensics via methods such as principal component and cluster analysis) to optimize conceptual site models, additional assessment, and remediation, thereby expediting the regulatory closure process;
    - Assisting with data migrations by efficiently cleaning and transforming data from disparate sources, including pdf reports and hand-written field notes; and
    - Reviewing data for anomalies quickly.

Additional Information

This proposal supports Pillar 4: Make the United States the Artificial Intelligence Capital of the World from the U.S. Environmental Protection Agency's (EPA) Powering the Great American Comeback Initiative. It also aligns with recent executive orders related to AI, which aim to streamline environmental permitting processes and enhance the adoption of AI within federal agencies to increase efficiency while complying with modern privacy standards.

Meeting notes from the authors' first brainstorm call were transcribed by Microsoft Copilot (a Generative AI assistant developed by Microsoft, released in 2023). These meeting notes were then uploaded to Copilot along with the ITRC proposal template to generate an initial draft of this proposal. This initial draft was then heavily reworked by the authors to attain a consensus proposal. Microsoft Copilot was occasionally used during the revision process to brainstorm how to reword specific phrases or sentences. This is an example of a "human-in-the-loop" approach to collaborating with AI and documenting the use of AI transparently.

The team would be composed of any interested academia and regulatory and regulated entities, including state and federal agencies, tribes, consultants, and industry. Many of the individuals listed below were active participants in the ITRC EDM team. The individuals listed below were actively involved in the development of this proposal, and the individuals and groups listed below have expressed interest in participating on the team:

State Agencies:
NC Department of Environmental Quality, Brian Pointer

Local Agencies:
City of Seattle, Jennifer Arthur, Ph.D.

Federal Agencies:

Federally Funded Research and Development Centers:
Pacific Northwest National Laboratory, Inci Demirkanli, Ph.D.

Academia:
Carnegie Mellon University, Ryan Sullivan, Ph.D.

Consultants:
Earthsoft, Inc., Dan McCarthy
Environmental Resources Management, Inc. (ERM), Meghan Eschbaugh, P.E.
Environmental Resources Management, Inc. (ERM), Theresa Kennedy
Environmental Works, Inc., Steve Brauner, Ph.D., P.E.
Geosyntec, Samantha Fuchs, Ph.D.
Geosyntec, Raphael Siebenmann
GES, David Cleland, PG, GISP
Gradient, Matthew Mayo, GISP, CPG, PG
GZA GeoEnvironmental, Austin Elliott
Hamp, Mathews & Associates, Inc., Stephen Zayko, P.E., CPG
Ramboll, Mary L. Cottingham
TRC, Kristen Brown
TRC, Chris Mickle, PMP
Verdantas, Jamie Combes
Verdantas, Tim Negley
Woodard & Curran, Inc., Sercan Ceyhan, Ph.D.
Woodard & Curran, Inc., Katie Elich
Woodard & Curran, Inc., Edward Piechowicz
Woodard & Curran, Inc., Victoria "Tori" Ward
Woodard & Curran, Inc., Shana Whitney

Professional Organizations:
International Commission on Environmental Data Management (ICEDM)