AI2ES Strategic and Implementation Plan

Make sure you are on slack!

https://join.slack.com/t/ai2es/shared_invite/zt-kbmbz1oh-ZGJjxr2GywXViqgmWo9zGg

Polls go to #general channel today

Link to the full plan

https://docs.google.com/document/d/1ntE79k8WoqIijxwxH2x6MGN3NRrG465VU0RnKmtihBw/edit?usp=sharing

Please make any changes as comments or suggestions!

AI2ES Vision Statement

The vision of the NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES) is to **create trustworthy AI methods that will be utilized by diverse ES user groups to revolutionize our understanding and prediction of high-impact atmospheric and ocean science phenomena and to create new educational pathways to develop a larger and more diverse AI/ES workforce**.

Strategic and Implementation Plan

- Transform our high-level goals from the proposal into actions
- Relatively detailed milestones for year 1, fewer details for years 2-5
- High-level science goals
- Lots of SMART goals
- Today's goal: familiarize yourself with the plan and make sure your name is on the teams you are in!

Overall science goals

- Creation of new trustworthy AI methods
- Discovery of novel scientific theories
- Improved understanding about fundamental environmental science processes
- Novel approaches to understanding trust between humans and AI methods in a variety of situations and for multiple types of end-users
- Creation of long-term educational pathways that will improve the diversity of the AI and ES workforce

Focus 1: Foundational AI Research

- Leaders: McGovern (OU), Gagne (NCAR), Ebert-Uphoff (CSU)
- Goal 1a: Develop explainable AI methods aligned with ES domain perspectives and priorities.
- Goal 1b: Develop physically based AI techniques for ES domains.
- Goal 1c: Develop robust AI prediction techniques, and empirically and theoretically validate their performance with adversarial data (e.g., missing data or intentionally wrong data).

Goal 1a: XAI

- Leaders: McGovern (OU), Gagne (NCAR), Ebert-Uphoff (CSU)
- Goal 1a.1: Develop XAI methods for ES data (including regression-based predictions, data with high spatiotemporal autocorrelations, and fielded data)
- Goal 1a.2: Develop XAI methods that integrate physics into the explanations (cross-cutting with Goal 1b)
- Goal 1a.3: Develop XAI methods that explain model failures
- Goal 1a.4: Develop XAI methods that facilitate knowledge and hypothesis discovery
- Goal 1a.5: Develop XAI approaches to effectively communicate (measured through RC research) estimated uncertainty to the end user and tailor these to the needs of the end-user
- Join us: reply to the polly!

Goal 1b: Physics-based AI

- Leaders: McGovern (OU), Hickey (Google) Gagne (NCAR), Ebert-Uphoff (CSU)
- Goal 1b.1: Develop techniques to ensure an AI method cannot fail in a physically implausible way
- Goal 1b.2: Develop physics-guided approaches to autonomous feature discovery
- Goal 1b.3: Develop hybrid models that incorporate physics-based AI
- Join us! Reply to the polly!

Goal 1c: Robust AI

- Leaders: McGovern (OU), Gagne (NCAR), Diochnos (OU)
- Goal 1c.1: Develop methods to allow transfer learning to be used to train AI models for rare weather phenomena
- Goal 1c.2: Develop robust semi-supervised and unsupervised learning algorithms for situations where reliable labels are not available
- Goal 1c.3: Develop theoretical and practical bounds on the robustness of the AI methods given class imbalance, a lack of reliable labels, and for adversarial situations (e.g. data may be missing or corrupted based on weather conditions)
- Join us! Reply to the polly!

Metrics and targets

- Novel AI methods/algorithms created and demonstrated to be useful for trustworthiness
  - Target: We do not set a target for the number of new algorithms here, as a meaningful metric is not the number of algorithms we create, but their impact on advancing our use cases. Specific targets for advancing our use cases are discussed below in Focus 2.
- Publications in top AI conferences, workshops, and journals (see Award Specific Conditions for a list of how we determine which AI/CS conferences to publish in)
  - Target: Submit at least 5 papers at top AI conferences, AI workshops, or AI-related journals each year after Y1.

Metrics and targets

- Interdisciplinary collaborations (as measured on student advising, publications, and presentations)
  - Target: At least 90% of all papers and presentations include collaborative authors and 100% of students include collaborative members of their committee. Note collaborative is defined as crossing institutions, sectors, and/or research areas.
- Presentations (including invited talks) in top conferences
  - Target: Present at least 10 talks in top conferences each year after Y1 (these can overlap with Focus 2 and 3)
  - Target: At least 2 invited talks per year across conferences

Metrics and targets

- Diversity of applications where the methods are used (e.g. ES use-cases, Research to Operations through our private industry and government collaborations, any additional uptake of the methods)
  - Target: Develop trustworthy for all use-cases by Y5 to inform users' decisions. Transition at least one method per use case to stakeholders by Y5.

Environmental Science use cases

- Convective
  - Leaders: Snook (OU), Gagne (NCAR), McGovern (OU), Demuth (NCAR)
  - Goal UC1a: Develop physically-based, robust AI methods that improve convective hazard prediction
  - Goal UC1b: Develop and use XAI methods to improve scientific understanding of convective hazards
  - Goal UC1c: Develop physically-based, robust AI and XAI methods that are deemed trustworthy by forecasters for convective hazard prediction
- Winter weather
  - Leaders: Thorncroft (Albany), Sulia (Albany), Bostrom (UW), Fagg (OU)
- Tropical Cyclones
  - Leaders: Ebert-Uphoff (CSU), Thorncroft (Albany), Gagne (NCAR), Demuth (NCAR)
- S2S
  - Leaders: Ebert-Uphoff (CSU), Barnes (CSU), Tissot (TAMUCC)
- Coastal Oceanography
  - Leaders: He (NCSU), Tissot (TAMUCC), Demuth (NCAR), Bostrom (UW), Williams (IBM), McGovern (OU)

Education, Workforce Development, Broadening Participation

- Goal 4a: Build a diverse AI for ES workforce by implementing and evaluating an applied AI for ES curriculum at the Community College Level and pilot test at an HSI/MSI.
  - Leaders: Davis (DMC), Tissot (TAMUCC)
- Goal 4b: Develop and share AI/ES curriculum for K-20 and workforce retraining.
  - Leaders: Rogers (CSU), Gagne (NCAR), McGovern (OU)
- Goal 4c: Broaden and train AI for ES through targeted internships, mentoring and recruitment.
  - Leaders: Hickey (Google), Williams (IBM), King (TAMUCC), Betz (DMC)

Knowledge Transfer

- Hickey (Google) and Neeman (OU) for data working group
- Gagne (NCAR) coding practices working group
- Williams (IBM), Hickey (Google), Griffin (Disaster Tech), Boukabara (NOAA) for R2O

AI2ES Code of Conduct

The AI2ES code of conduct was derived from the codes of conduct from the American Meteorological Society, the American Geophysical Union, and the American Association of Artificial Intelligence.

1. Members shall carry out their activities with integrity and the highest ethical standards. Members shall adhere to the AI2ES ethical code of conduct.
2. Members will create and uphold a safe, open, welcoming, inclusive, and professional environment for learning, conducting, and communicating science with integrity, respect, fairness, trustworthiness, and transparency. This includes learning and using good practices for intercultural and diverse collaborations, maintaining the dignity of all individuals, and valuing diversity in all forms.
3. Members must refrain from all forms of discrimination, harassment, and bullying.
4. In addition to their own conduct, Members who either witness or are made aware of discrimination, harassment, or bullying by others should actively seek to prevent, report, and/or otherwise mitigate the offensive behavior.
5. Members are also expected to adhere to their employer's code of conduct and ethics.
6. Violations of the code of conduct can be reported to either anyone in the leadership team or anyone in the EAB (if the reporter is not comfortable reporting to the leadership team for any reason). Any reports will be kept confidential and addressed carefully and confidentially. If a member is violating the code of conduct or ethics, the senior leadership team will determine what sanctions need to be imposed. This can include requiring DEI training or even banning them from AI2ES for serious offenses. Serious offenses may be reported to the individual's institution for further action.

AI2ES Code of Ethics

The AI2ES code of ethics was derived from the codes of ethics from the American Meteorological Society, the American Geophysical Union, the American Association of Artificial Intelligence and Google's AI Principles.

1. Integrity: Members will act with honesty and integrity in the interest of the advancement of science, take full responsibility for the trustworthiness of their research and its dissemination, and treat others with courtesy, equity, and fairness. Members shall not commit scientific misconduct, defined as fabrication, falsification, or plagiarism. Scientific error or incorrect interpretation of research data that may occur as part of the scientific process does not constitute scientific misconduct.

AI2ES Code of Ethics

2. Publications and Research:
    a. Members will employ research methods to the best of their understanding and ability, base conclusions on critical analysis of the evidence, and report findings and interpretations fully, accurately, and objectively, including characterization of uncertainties.
    b. Members shall be fair, impartial, objective, and prompt peer reviewers, maintaining confidentiality when requested.
    c. Members will respect the AI2ES intellectual property agreement.
    d. Members will share data and findings openly and promptly. Members will maintain clear, accurate records of research in ways that will allow verification and replication of their work by others. Members will follow best practices for data management, accessibility, and preservation.
    e. Members will promptly correct the literature when errors in their own work are detected.
    f. Members will take responsibility for the integrity of their contributions to all publications, funding applications, reports, and other representations of their research. Author credit should be given to those who have made meaningful contributions to publications and members will acknowledge the names and roles of those who made significant contributions (such as ideas and scientific discussion) to the research.
    g. Members will act or intercede where possible to prevent scientific misconduct and to report as necessary to AI2ES and any relevant professional societies.

AI2ES Code of Ethics

3. Stewardship of the Earth: Members have an ethical obligation to weigh the societal benefits of their research against the costs and risks to human and animal welfare, heritage sites, or other potential impacts on the environment and society. Members need to be aware of legal requirements in this area.
4. Public Communication:
    a. Members, when representing AI2ES, will limit professional comments to their areas of scholarly expertise when engaged in public discussions about the application and importance of research findings and will clearly distinguish professional comments from their opinions based on personal views.
    b. Members have an ethical obligation to responsibly, accurately, and clearly inform the public about natural resources, hazards, and other geoscience phenomena of importance to the well-being of Earth and society.
    c. Members have an ethical obligation to foster public awareness and understanding of AI, computing, related technologies, and their consequences.

AI2ES Code of Ethics

5. When creating AI systems, members will:
    a. Ensure that the public good is the central concern during all professional computing work
    b. Give comprehensive and thorough evaluations of AI2ES AI algorithms and their impacts, including analysis of possible risks.
    c. Recognize and take special care of AI systems that become integrated into the infrastructure of society.
6. Members will create AI systems that will:
    a. Avoid harm
    b. Protect the Earth and its environment including human and animal welfare.
    c. Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.
    d. Be fair and take action not to discriminate.
    e. Respect privacy.
    f. Honor confidentiality.
    g. Avoid creating or reinforcing unfair bias.
    h. Uphold high standards of scientific excellence.

AI2ES Code of Ethics

7. Conflict of Interest: Members shall disclose all relevant relationships, financial or otherwise, that might be perceived to unduly influence the outcome of their research.
8. Reporting: Violations of the code of ethics can be reported to either anyone in the leadership team or anyone in the EAB (if the reporter is not comfortable reporting to the leadership team for any reason). Any reports will be kept confidential and addressed carefully and confidentially. If a member is violating the code of conduct or ethics, the senior leadership team will determine what sanctions need to be imposed. This can include requiring additional training or even banning them from AI2ES for serious offenses. Serious offenses may be reported to the individual's institution for further action.