**Review on Responsible AI for conservation**

*( Wearn, O. R., Freeman, R., & Jacoby, D. M. (2019). Responsible AI for conservation. Nature Machine)*

1. First of all, it is obvious that AI(Artificial Intelligence) and ML(Machine Learning) are used in many fields for protecting the nature:
   1. predicting the extinction risk of thousands of species;
   2. assessing the global footprint of fisheries;
   3. and identifying animals and humans in wildlife sensor data recorded in the field.

Commercial companies supports:

1. Microsoft’s AI for Earth
2. Google’s AI for Social Good
3. However, their( AI and ML) misuse could have severe real-world consequences for people and wildlife. The opaque nature of some ML algorithms means that the potential for unintended consequences may be high and this could have real-world consequences for people and wildlife:
4. it can be difficult to identify the implicit assumptions of an algorithm (for example, how much of the contextual background information it is using when identifying species in images).
5. it might be unclear when an algorithm is being asked to make predictions beyond the scope of the training data.
6. an algorithm might not be easily interrogated as to why it made a particular decision.
7. Better metrics are needed, since simple accuracy metrics are unlikely to provide a good indicator of success when an algorithm is transferred to new datasets.
8. Better ethical oversight of the use of AI in conservation is needed.
9. Two potential goals for the conservation and AI communities to tackle in the immediate term:
   1. the development of metrics to better allow conservationists to assess the usefulness of an algorithm;
   2. and the formulation of ethical guidelines for the responsible use of AI in conservation.

**Review on Five Ps: Leverage Zones Towards Responsible AI**

Nabavi, E., & Browne, C. (2022). Five Ps: Leverage Zones Towards Responsible AI. arXiv preprint

1. What is Five Ps?

problem, parameter, process, pathway, and purpose

1. What can Five Ps do for responsible AI?
   1. Problems identified in the Parameter zone are tractable (modifiable, mechanistic) characteristics of an AI system that are commonly targeted by AI developers to improve the responsibility of AI. They are typically smaller visible flaws that are usually addressed through engineering solutions such as tweaking algorithms and parameters. The effort to fix these is small, and changes in this zone are incremental and may have a negligible effect on the problem’s underlying structure or dynamics. They are important markers of the problem, but they are often symptomatic and not the root cause of the problem.
   2. Problems identified in the Process zone consider the wide range of interactions between the feedback elements of an AI system that drive the internal dynamics, including social and technical processes associated with how the AI is designed, built, and deployed. This might include activities that speed up development times, or actively responding to emerging trends in the data. Changes in this zone are likely to result in resolving issues as they emerge or amplifying the effect of assumptions.
   3. Problems identified in the Pathway zone consider the ways through which information flows, the rules are set, and the power is organized. For example, improving transparency of how algorithms are employed, the governance or legislation of their use, or putting the ownership of data back into the consumer’s hands. These changes are structural to the system that allows the AI to operate, and result in establishing new patterns of behavior and agency.
   4. Issues identified in the Purpose zone have the most potential to affect change in a system. These relate to the norms, values, goals, and worldviews of AI developers that are embodied in the system. It includes the underpinning paradigms based on which the system is imagined, and the ability to transform entirely and imagine new paradigms. Framing perceived problems in this zone serves to act as a compass to guide the developers to align with the fundamental purpose of the system.