Some stuff about AI for Earth

Dan Morris
Microsoft AI for Earth
dan@microsoft.com
aka.ms/aiforearth
aka.ms/dmorris

(pro tip: pictures of animals put your audience in a good mood)

(it’s working, right? aren’t you in a good mood already?)

gratuitous pictures of animals that aren’t related to my talk
Question #1, as posed by the panel organizers…

“Please describe your organization or project and your role in it.”

Can do!!!
AI for Earth is three things…

Grants
aka.ms/ai4egrants

Data
aka.ms/ai4edata

I’m keeping you in suspense about the third thing we do; I’ll tell you later, I promise.
AI for Earth: grantee stories

Microsoft AI for Earth grantee gallery

Profiles

Microsoft AI for Earth grantee gallery

Showcasing the work of Microsoft AI for Earth grant recipients

Profiles

AI for Earth supports organizations all around the world that are working on challenges in biodiversity conservation, climate change, agriculture, and water. Read more about the AI for Earth grantees and the amazing projects they’re working on in environmental sustainability.

Collapse

AdaViv

Using AI to unleash the potential of urban agriculture

AdaViv is developing an adaptive and efficient indoor growing system on Azure that uses sensors, actuators, and machine learning to monitor plant growth, predict yields, detect diseases, and understand precisely how nutrients, environment, and light are affecting plant growth. This system will help indoor producers attain higher yields, precise quality control, and hyper-efficient production.

Ag-Analytics

Improving agriculture forecasting and conservation practices

aka.ms/ai4egrantees
AI for Earth: big data

NEXRAD
Radar data from 159 stations across the U.S.
NEXRAD on Azure Open Datasets

Daymet
Gridded temperature data across North America
Daymet on Azure Open Datasets

NAIP
High-resolution aerial imagery across the continental U.S.
NAIP on Azure Open Datasets

MODIS surface reflectance
500m-resolution global daily surface reflectance dating back to 2000
MODIS on Azure Open Datasets

NOAA Global Forecast System
15-day U.S. hourly weather forecast data
GFS on Azure Open Datasets

NOAA Integrated Surface Data
Worldwide hourly weather history data
ISD on Azure Open Datasets

NASADEM
Global topographic survey
NASADEM on Azure Open Datasets

Harmonized Landsat Sentinel-2
30m-resolution satellite imagery for North America
HLS on Azure Open Datasets

aka.ms/ai4edata
Big data: you can help us help you!

We want to know about…

• The public data sets you use
• The formats that delight/annoy you
• The tools you use to access/process data
• Your experiences with the data we host
• Your experiences with the data other clouds host

aiforearthdatasets@microsoft.com
And let’s make some noise for “small data”!

http://lila.science
Question #2, as posed by the panel organizers…

“How does your group use ML?”

Let me tell you all about it!
AI for Earth is three things (for real now)…

Grants
aka.ms/ai4egrants

Data
aka.ms/ai4edata

ML stuff
aka.ms/ai4ecode
AI for Earth data science:
Using ML to help conservation scientists spend less time clicking stuff, and more time doing conservation.

Accelerating *camera trap* surveys
github.com/microsoft/cameratraps

Accelerating *acoustic* wildlife surveys
github.com/microsoft/belugasounds
github.com/microsoft/multi_species_bioacoustic_classification

Accelerating *aerial* wildlife surveys
github.com/microsoft/aerial_wildlife_detection
github.com/microsoft/arcticspeaks

Accelerating *land cover* surveys
aka.ms/landcovermapping

All the fun animal-related work in one place
aka.ms/biodiversitysurveys
A couple more things we do to help move the needle on ML for conservation and sustainability…
Helping to get *labeled* data out there for areas we work in…

http://lila.science
Camera trap data…

Snapshot Karoo

Overview
This data set contains 14889 sequences of camera trap images, totaling 38074 images, from the Snapshot Karoo project, part of the Snapshot Safari network. Using the same camera trapping protocols at every site, Snapshot Safari members are collecting standardized data from many protected areas in Africa, which allows for cross-site comparisons to assess the efficacy of conservation and restoration programs.

Karoo National Park, located in the arid Nama Karoo biome of South Africa, is defined by its endemic vegetation and mountain landscapes. Its unique topographical gradient has led to a surprising amount of biodiversity, with 58 mammals and more than 200 bird species recorded, as well as a multitude of reptilian species.

Labels are provided for 38 categories, primarily at the species level (for example, the most common labels are gemsbokoryx, hartebeestred, and kudu). Approximately 85.02% of images are labeled as empty. A full list of species and associated image counts is available here.
Land cover data…

Chesapeake Land Cover

Overview

This dataset contains high-resolution aerial imagery from the USDA NAIP program [1], high-resolution land cover labels from the Chesapeake Conservancy [2], low-resolution land cover labels from the USGS NLCD 2011 dataset [3], low-resolution multi-spectral imagery from Landsat 8 [4], and high-resolution building footprint masks from Microsoft Bing [5], formatted to accelerate machine learning research into land cover mapping. The Chesapeake Conservancy spent over 10 months and $1.3 million creating a consistent six-class land cover dataset covering the Chesapeake Bay watershed. While the purpose of the mapping effort by the Chesapeake Conservancy was to create land cover data to be used in conservation efforts, the same data can be used to train machine learning models that can be applied over even wider areas.
Aerial survey data...

NOAA Arctic Seals

This data set contains about one million thermal/RGB image pairs, representing a 2016 aerial survey of sea ice habitat in U.S. waters of the Chukchi Sea, conducted by NOAA fisheries. Annotations indicate the locations of approximately 7000 seals in these images. This data set is provided to encourage the machine learning community to advance the state of the art in detection with an extremely imbalanced data set (the vast majority of images are empty), image registration (thermal and RGB images are not perfectly aligned), and multimodal fusion in detection.
Helping to advance the state of the art for competitions …

DrivenData and Microsoft AI for Earth are launching an exciting new kind of challenge with a $20,000 prize pool. Sign up to get notified!

We'll email you when the competition goes live. That's it.
Some stuff about AI for Earth

Dan Morris
Microsoft AI for Earth
dan@microsoft.com
aka.ms/aiforearth
aka.ms/dmorris