Garden: A FAIR Framework for Publishing and Applying AI Models

for Translational Research in Science, Engineering, Education, and Industry

Aristana Scourtas (aristana@uchicago.edu), Ben Blaiszik (blaiszik@uchicago.edu), Ryan Chard, ZhuoZhao Li, Logan Ward, Tyler Skluzacek, Ben Galewsky, KJ Schmidt, Nathan Pruyne, Jim James, Isaac Darling, Braeden Cullen, Aadit Ambadkar, Ethan Truelove, Marcus Schwarting, Eliu Huerta, Rebecca Willett, Rafa Gomez-Bombarelli, Dane Morgan, Ian Foster (foster@anl.gov)

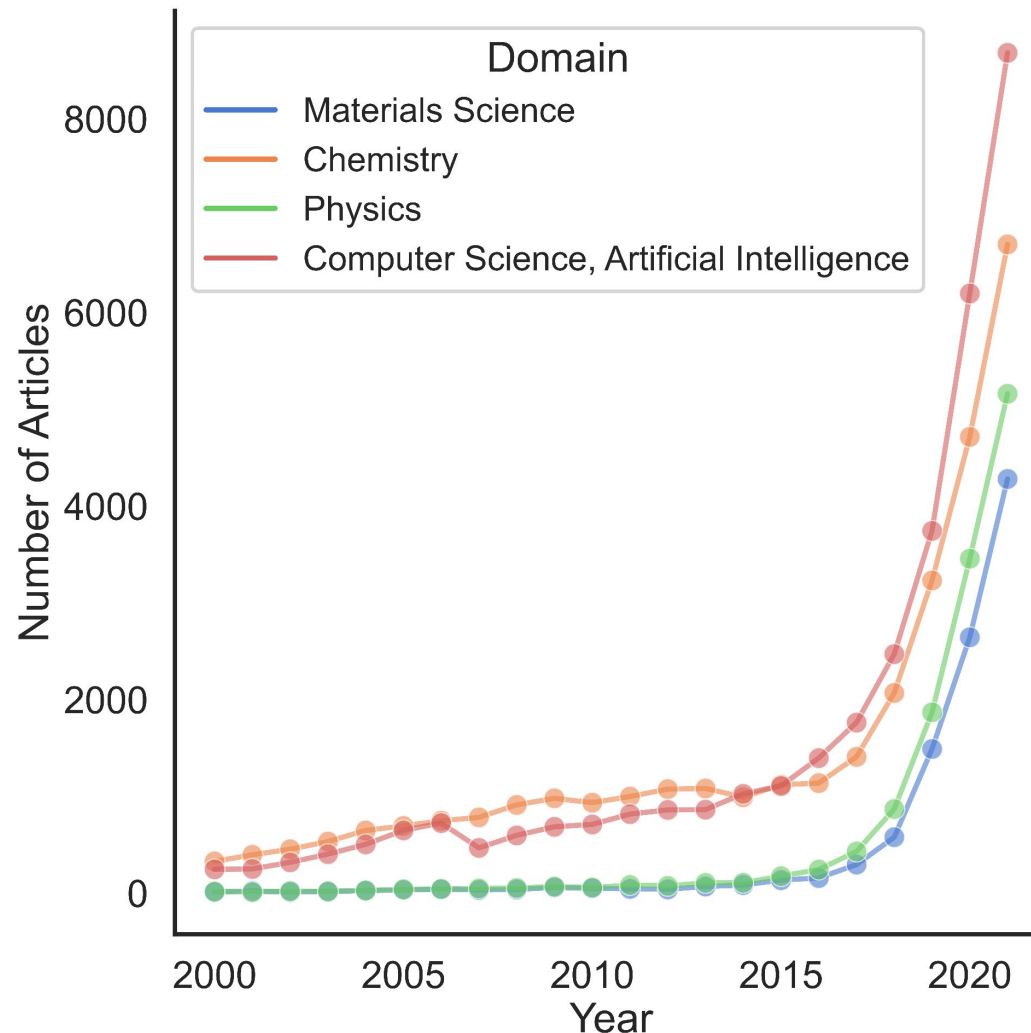**https://www.dlhub.org**

The Growing Importance of ML and Data in the Sciences

Data and ML are becoming key drivers of scientific progress

Methods and Data:
https://github.com/blaiszik/ml_publication_charts

How do we use these models?

For a given study:

- Where is the code?
- Where are the trained models?
- Where are the training data?
- How can I reproduce these results?

Without all of these pieces, progress is drastically slowed

Need models and data to be FAIR:

Findable

Accessible

Interoperable

Reusable

Location of many ML models after a paper is finished

DLHub for FAIR Models

A simple way to find, share, publish, and run machine learning models

1 ML Model Submission
- Collect ML models in

2 Container Creation
- Create portable containers of models and register in **funcX**

3 Model Catalog
- Create a searchable index of data

REST API

Model Publisher
- DOI for citing model
- Landing page
- Run in the cloud
- API to model!

Model Consumer

REST API
- Find
- Run (in cloud)

4
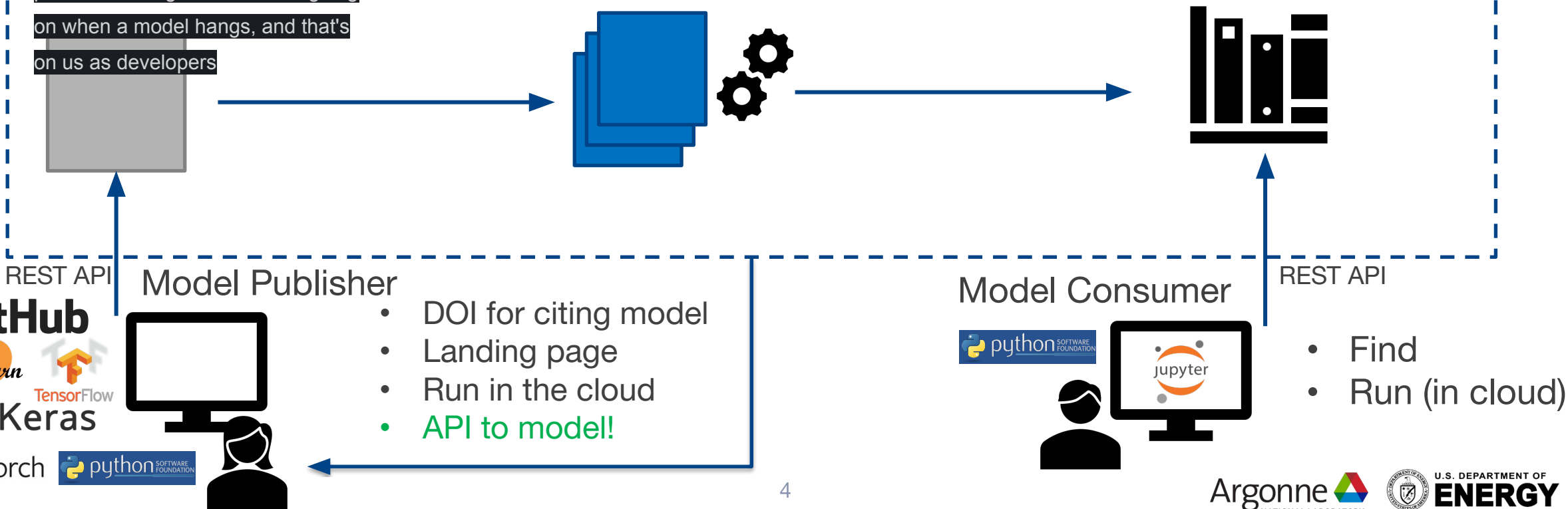
DLHub Containers with funcX

Container

```
dlhub.json
Dockerfile
maskrcnn_model
model_final.pth
requirements.txt
runtime.txt
```

user-specified

funcx_endpoint
dlhub_sdk
home_run
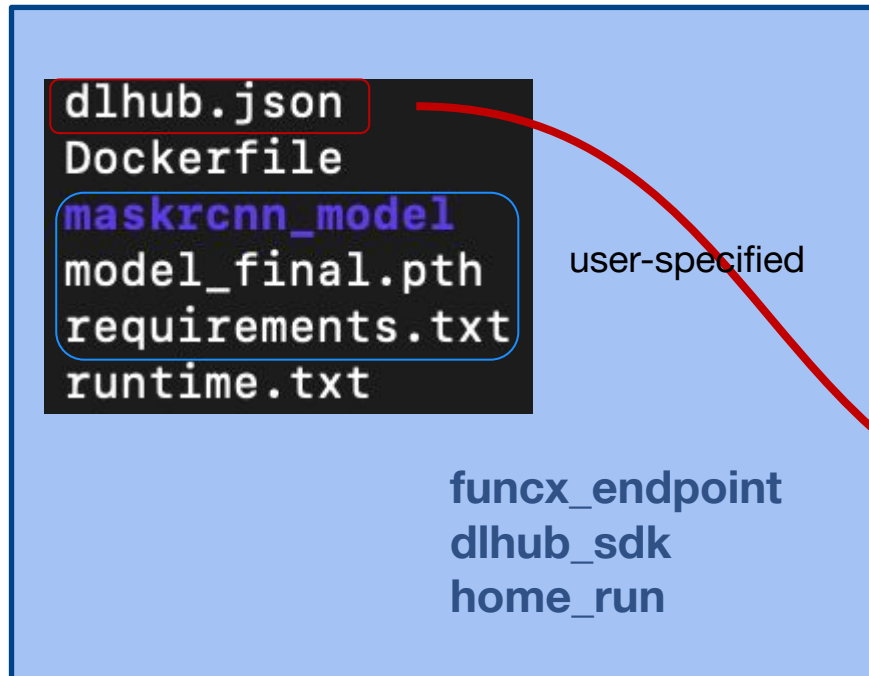
we then register the container and the function `dlhub_run()` with funcX

dlhub_run(event)

```
from home_run import create_servable
with open("dlhub.json") as fp:
    shim = create_servable(json.load(fp))
```

dlhub.json contains all servable-specific info

Garden for FAIR Models and Data

An ecosystem to connect researchers, provide resources, and validate models for the community

Garden for FAIR Models and Data

Our goal is to make it easy for the user to assess applicability, compatibility, and reliability

Then, the user can take action:
- Cite/share
- Discover
- Understand
- Decide

Black Hole Merger Event Detection

GPU-enabled

CPU-enabled

Classification

E. A. Huerta, Asad Khan, Xiaobo Huang, Minyang Tian, Maksim Levental, Ryan Chard, Wei Wei, Maeve Heflin, Daniel S. Katz, Volodymyr Kindratenko, Dawei Mu, Ben Blaiszik & Ian Foster

FAIR Score

9 /10

DOI: 10.26311/zwfw-3n68 [Get Citation] **v: 1.2** ☆ Add Favorite

100,563 invocations
☆ 1k favorites

Here, we develop an AI model for the detection of black hole merging events from LIGO-observed spectral image data, trained on the Hardware-Accelerated Learning (HAL) cluster.

📄 Paper   📦 Container image
 Code   Datasets

Software Testing   1 test failing - details

- ● Python 3.9 - GPU   2 days ago
- ● Python 3.9 - CPU   2 days ago
- ● Python 3.8 - GPU   2 days ago

Reliability Testing   Accuracy

- ● LIGO Validation Set 1   0.95   5 minutes ago
- ● LIGO Training Set   1.00   5 minutes ago
- ● LIGO Validation Set 2   1.00   1 day ago

Performance Testing

CPU 20 iterations/s   **GPU** 100 iterations/s   **Hardware Accelerator** 500 iterations/s

Model Playground

[Open in WholeTale]

[Open in Hugging Face]

Use the Model

```
from garden_sdk import GardenClient
gc = GardenClient()

id = "10.18126/sq79-ki6m"
gc.run(id, my_data)
```
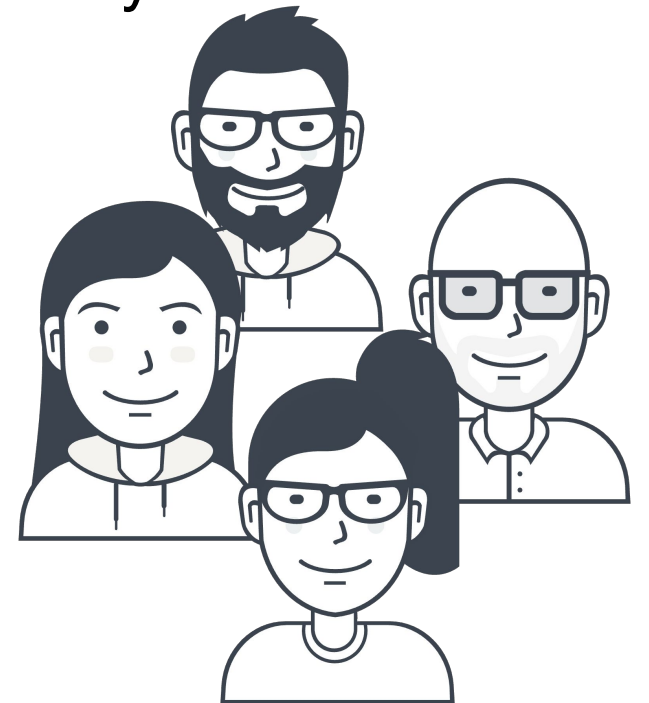
We're hiring 🌱

Interested in building Garden with us, or know someone who is?
We're currently hiring for software developers to help us deliver
FAIR infrastructure and models to the community

Reach out if you're interested in learning more!

Contact: Aristana Scourtas (aristana@uchicago.edu)
KJ Schmidt (kjschmidt@uchicago.edu)
Ben Blaiszik (blaiszik@uchicago.edu)

Thank You!

Garden

DLHub

https://www.dlhub.org

Funding: 2018 Argonne Adv. Computing LDRD

FOUNDRY
DATA, MODELS, SCIENCE

https://www.foundry-ml.org

MATERIALS DATA FACILITY

https://www.materialsdatafacility.org

Contact: Aristana Scourtas (aristana@uchicago.edu)
KJ Schmidt (kjschmidt@uchicago.edu)
Ben Blaiszik (blaiszik@uchicago.edu)