End-to-end acceleration of machine learning in gravitational wave physics

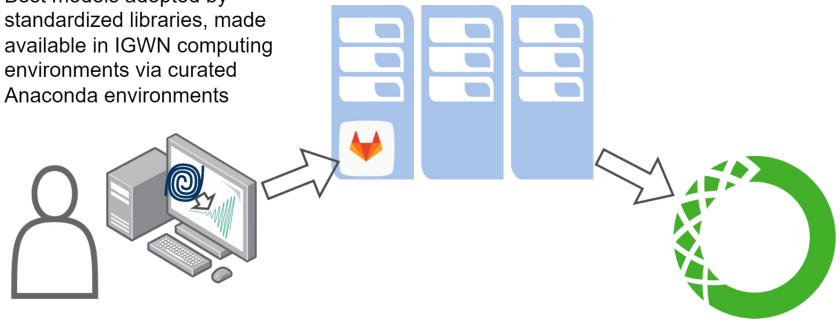
Alec Gunny[1,*], Ethan Marx[1], William Benoit[2], Saleem Muhammed[2], Ryan Raikman[1], Deep Chatterjee[1], Eric Moreno[1], Dylan Rankin[1], Michael Coughlin[2], Philip Harris[1], Erik Katsavounidis[1]

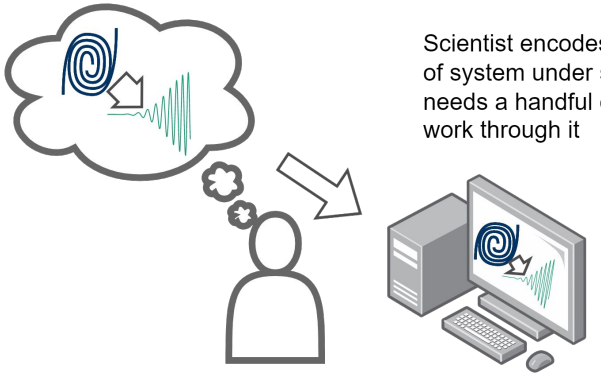
Fast Machine Learning for Science Workshop 2022
10/5/2022

* - presenter
1 - Massachusetts Institute of Technology
2 - University of Minnesota

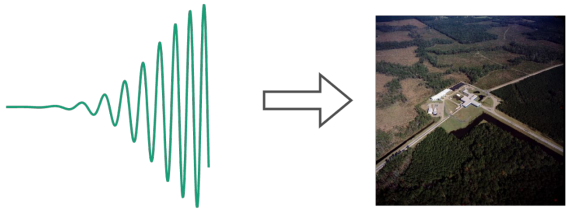
Traditional Gravitational Wave Physics Software Stack

Best models adopted by standardized libraries, made available in IGWN computing environments via curated Anaconda environments





Scientist encodes mental model of system under study, only needs a handful of examples to work through it



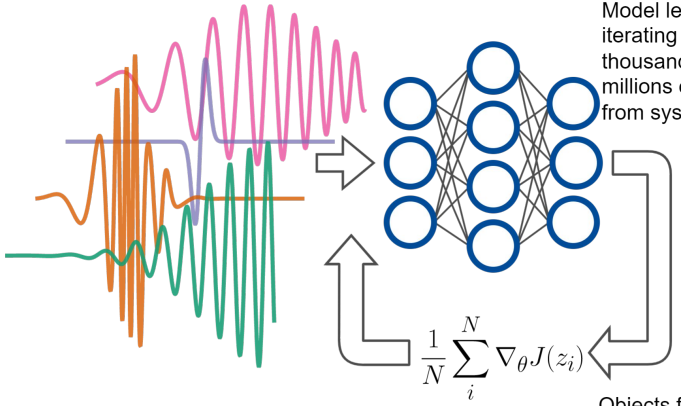

```python
class Waveform:
    def __init__(
        self,
        m1, m2,
        ra, dec, psi,
        geocent_time,
        **other_params
    ):
        ...
```

```python
class Detector:
    def __init__(
        self,
        vertex, azimuth,
        **other_params
    ):
        ...
    def project_waveform(
        self, waveform,
    ):
```

Object-oriented, intuitive workflows that map on to real physical concepts

Often designed for deep analysis of one waveform at a time

Traditional Machine Learning Software Stack



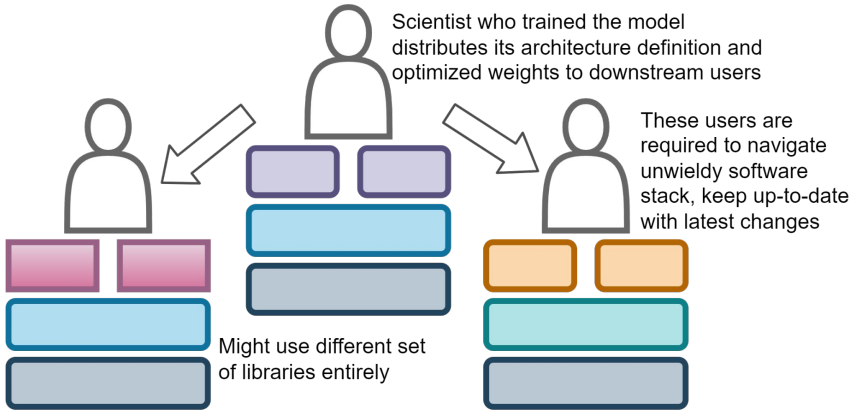
Model learns by iterating through thousands or millions of samples from system

$$\frac{1}{N}\sum_{i}^{N}\nabla_{\theta}J(z_i)$$

Objects framed in terms of tensors and operations between them

Take advantage of parallel processing of hardware accelerators e.g. GPUs

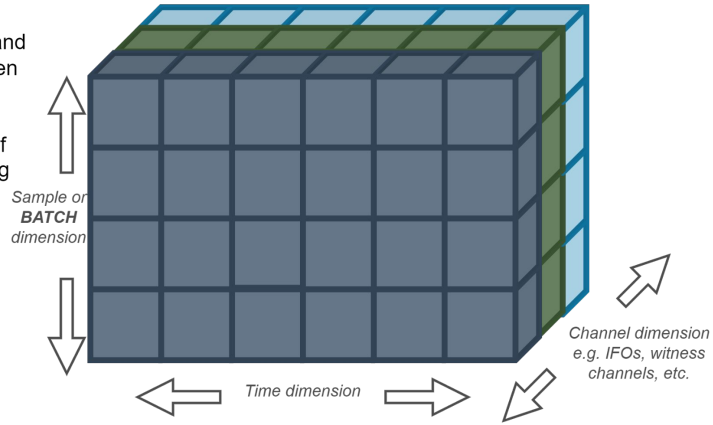
Use continuous functions to keep track of gradients automatically



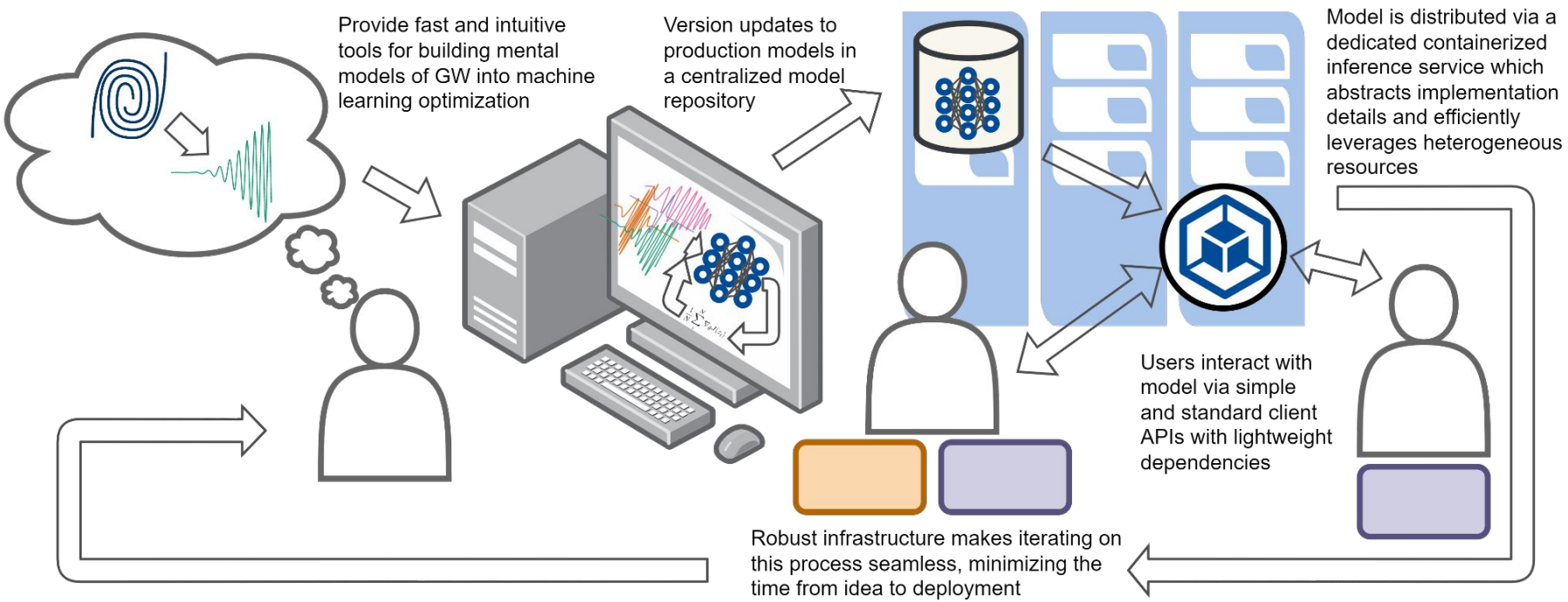
Scientist who trained the model distributes its architecture definition and optimized weights to downstream users

These users are required to navigate unwieldy software stack, keep up-to-date with latest changes

Might use different set of libraries entirely



ML4GW/HERMES - MLOps for fast end-to-end deployment



ML4GW

Training utilities for common GW operations, e.g. projecting raw GW waveforms to interferometer responses

HERMES

Inference-as–a-service deployment tools designed for streaming timeseries