

# Building Community Oriented AI Foundation Models in Heliophysics

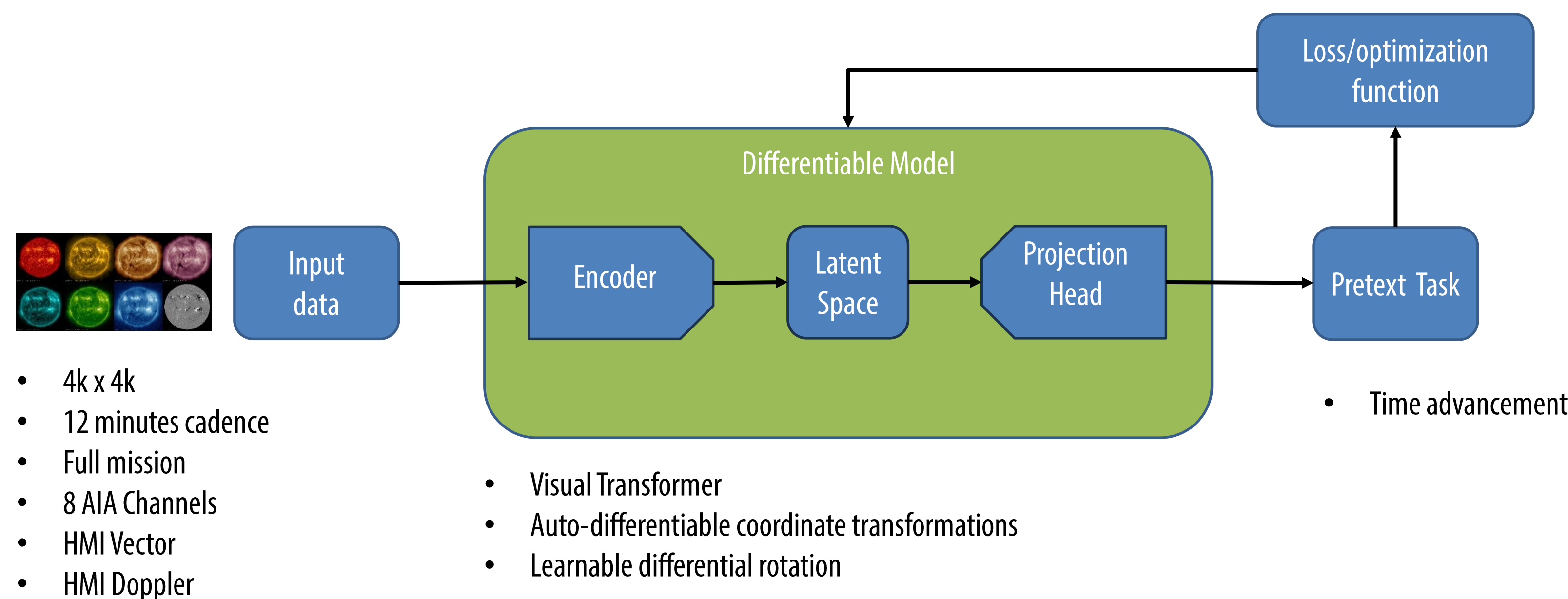
Andrés Muñoz-Jaramillo, Madhulika Guhathakurta and HelioFM team

- AI Foundation models (FMs) are models that are trained on vast datasets so that they can be applied across a wide range of use cases.
- We are building a Foundation Model using SDO data, a transformer architecture, and a range of downstream applications.
- For a scientific community, the value of a FM goes beyond the ability to train a model with lots of data: it's an opportunity to build a shared AI infrastructure.
- We are building auto-differentiable infrastructure that can be reused by other AI projects (data preprocessing, data-loaders, metrics, coordinate transformations, uncertainty estimation, baselines, benchmarks, uncertainty).
- We are looking for help shaping and make this model and shared ecosystem the best in can possibly be.

## Self-supervision setup

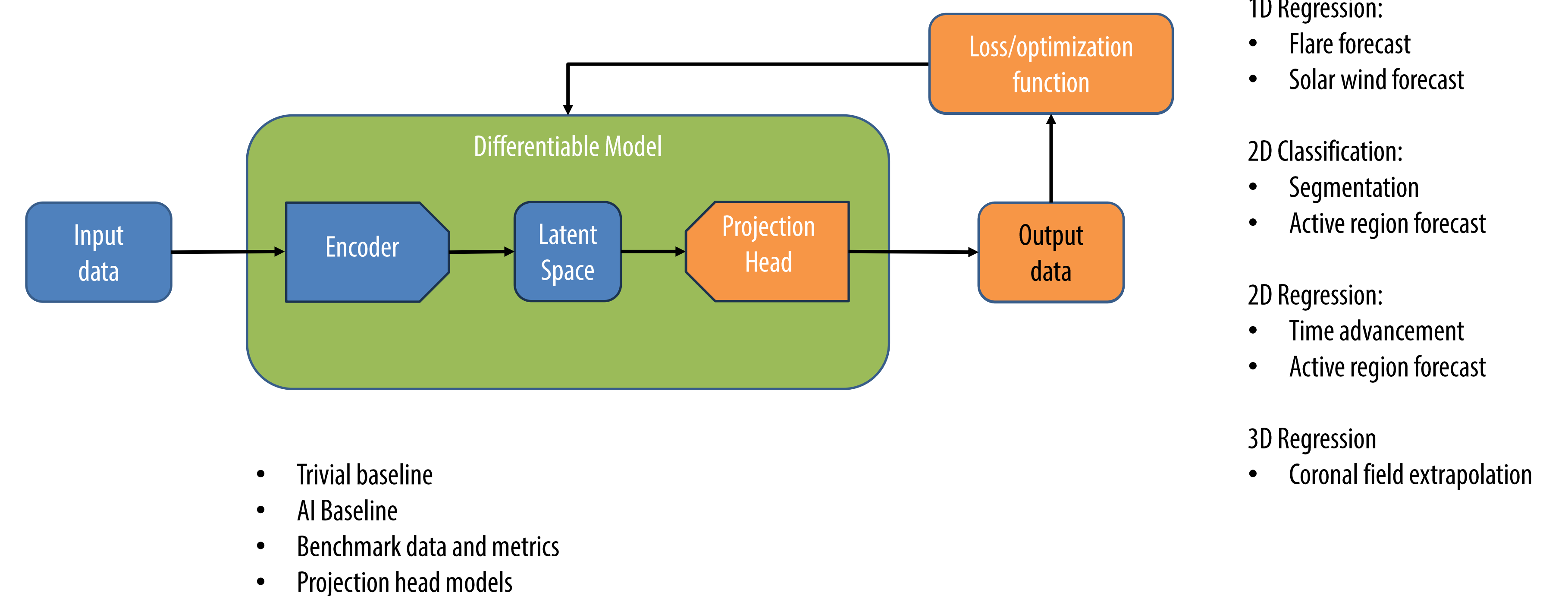
- Self-supervision allows training an encoder architecture on vast quantities of data without the need from human annotation.

- Mean Square Error
- Signed and unsigned flux (HMI)
- Differential Emission Measure



## Downstream Applications

- Once trained, the encoder can be frozen and only the projection head can be fine tuned to any particular application, without the need for the same volume of data.



## AI infrastructure for the Heliophysics community

- Our goal is to build common AI infrastructure that facilitates all aspects of the AI application lifecycle:



- Helio extension to IBM's TerraTorch framework (AI training and validation using a single config file).
- Preprocessing pipeline that builds on top of SDOML
- Auto-differentiable coordinate transformations.
- Auto-differentiable metrics suite that can be used as loss or diagnostic metric.
- Automatic ensembling.
- Zero-shot approximation to a Bayesian posterior.
- Packaging data and metrics to create benchmark standards.

## How can you help?

- Please contact us if you are interested in using the foundation model to train downstream applications.
- Please contact us if you are interesting in sharing and building AI infrastructure.
- Help us uncover technical challenges faced by the heliophysics community.



[bit.ly/heliofm-science-survey](https://bit.ly/heliofm-science-survey)

More info: [andres.munoz@swri.org](mailto:andres.munoz@swri.org)