# Evaluating the Software Carbon Intensity of Foundation Models

## Scope

Foundation models (e.g. BERT and GPT-3) refer to large scale AI models that are trained on massive amounts of data and can be adapted to a wide range of tasks (e.g. language, vision, human interaction etc.). Although foundation models are very powerful, they are extremely resource hungry with significant carbon emissions because of their scale. This project aims to create a comprehensive case study on evaluating the impact of foundation AI models on the environment by applying the Software Carbon Intensity (SCI) specification.

**Why does this project help the GSF meet its mission?**

GSF recently published the SCI specification, but the community still needs case studies to learn how to use SCI to evaluate the environmental impact of various software. This project will help GSF promote the SCI specification by providing a comprehensive case study on applying SCI to foundation AI models. The methodology we use and the tools we develop could help other GSF members to apply SCI to other AI models.

**Can you define the boundaries of this project so we can understand how it’s related to other GSF projects?**

The lifecycle of foundation models includes training, adaptation, and inference. This project will only focus on studying the environmental impact of foundation models at the inference stage.

## Team

Dr. Ziliang Zong, A Ph.D. student and an undergraduate student – Texas State University

A team of GSF ML experts to provide review capabilities

Supporting data providers:

Henry Richardson – Watttime

Benjamin Davy - Boavizta

## Structure

**How will the project be structured and managed?**

We expect the project to have three phases. The first phase will be primarily done by the Texas State Team to select several foundation models, write scripts to deploy these models, and develop tools to measure the energy consumption of these models. In the second phase, we will work with Watttime and Boavizta to report preliminary studies to the standard working groups and ask for internal feedback from other GSF members. In the third phase, we will make our study publicly available by writing a detailed report on the GSF website or publishing a paper.

**Has a Chair been identified?**

Ziliang Zong

**Is this project consensus based or maintainer based?**

It is consensus based.

**Will the team collaborate via GitHub or Google Docs?**

We will use Github for code sharing and Google Docs for discussions and progress reports.

**Which working group do you believe this project will fall under?**

Standard Working Group.

## Required Resources

**Can you list any resources you will need to complete this project, from GitHub repos, to people to money?**

Hardware: Texas State University will provide the required hardware to support his study.

Github: A Github repo for code sharing and discussions.

People time: Ziliang Zong (2 months), Ph.D. student (4 months), undergraduate student (3 months)

## Deliverables and Timeline

**Can you lay out what the key deliverables will be and a rough timeline of delivery?**

**Every deliverable will be reviewed by the Standards working group.**

//Timeline has been updated because the start date of this project was delayed for 1 month.

Deliverable 1 (by April 15): Several well-selected foundation models. These models may be revised based on currently available open-source models.

Deliverable 2 (by May 30): Power measurement tools for measuring the energy consumption of foundation AI models

Deliverable 3: (by June 30): Preliminary results of SCI scores of evaluated foundation models.

Deliverable 4: (by July 31): Technical report/paper of our study.

### Project Milestone dates

Milestone 1 (by April 15): Successful deployment of several foundation models.

Milestone 2 (by May 30): Accurately measure the real-time energy consumption of selected foundation models.

Milestone 3 (by June 30): Release the preliminary SCI scores of evaluated foundation models to GSF for internal discussions.

Milestone 4 (by July 31): Release the power measurement tools, selected foundation models, final SCI scores of different foundation models. Publish a technical report/paper.