THE UNIVERSITY OF CHICAGO

Self-Driving Systems for Physics White Paper

Cecilia Tosciri
on behalf of the group

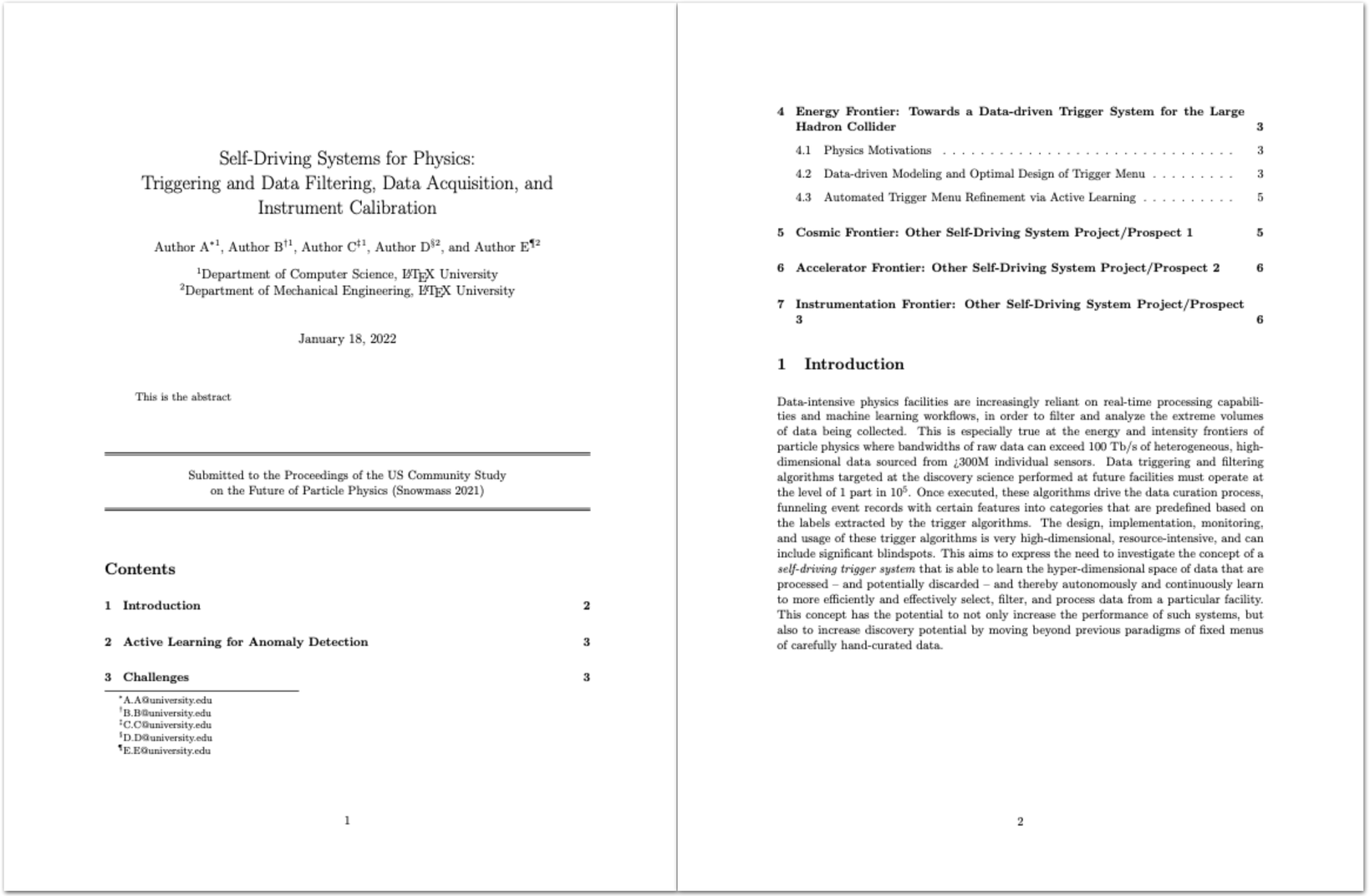
Comp F3 White Paper planning

19th January 2022

White Paper Planning

- Main goal: identify and discuss the relevance of dynamic feedback and data-driven approaches for physics systems (self-driving systems)
- Already significant attention from different perspectives (Energy, Cosmic, Accelerator Frontiers)
- Possible topics to be discussed: active learning for anomaly detection, reinforcement learning, interpretability, technical points and common challenges (algoriths and hardware), monitoring, current technologies
- Identified different works/projects that can be discussed as use-cases for the different approaches:
  - Energy Frontier: Self-Driving Trigger for the LHC (UChicago)
  - Cosmic Frontier: Cosmology telescope surveys (Brian Nord)
  - Accelerator Frontier: Accelerator controls for self-driving systems (Nhan Tran)
  - Instrumentation Frontier (?) -> contributions on this direction are more than welcome!

Link and contacts



Self-Driving Systems for Physics:
Triggering and Data Filtering, Data Acquisition, and Instrument Calibration

Author A[*1], Author B[†1], Author C[‡1], Author D[§2], and Author E[¶2]

[1]Department of Computer Science, LaTeX University
[2]Department of Mechanical Engineering, LaTeX University

January 18, 2022

This is the abstract

Submitted to the Proceedings of the US Community Study on the Future of Particle Physics (Snowmass 2021)

Contents

[*]A.A@university.edu
[†]B.B@university.edu
[‡]C.C@university.edu
[§]D.D@university.edu
[¶]E.E@university.edu

1 Introduction

Data-intensive physics facilities are increasingly reliant on real-time processing capabilities and machine learning workflows, in order to filter and analyze the extreme volumes of data being collected. This is especially true at the energy and intensity frontiers of particle physics where bandwidths of raw data can exceed 100 Tb/s of heterogeneous, high-dimensional data sourced from ¿300M individual sensors. Data triggering and filtering algorithms targeted at the discovery science performed at future facilities must operate at the level of 1 part in 10^5. Once executed, these algorithms drive the data curation process, funneling event records with certain features into categories that are predefined based on the labels extracted by the trigger algorithms. The design, implementation, monitoring, and usage of these trigger algorithms is very high-dimensional, resource-intensive, and can include significant blindspots. This aims to express the need to investigate the concept of a *self-driving trigger system* that is able to learn the hyper-dimensional space of data that are processed – and potentially discarded – and thereby autonomously and continuously learn to more efficiently and effectively select, filter, and process data from a particular facility. This concept has the potential to not only increase the performance of such systems, but also to increase discovery potential by moving beyond previous paradigms of fixed menus of carefully hand-curated data.

- A draft of the WP can be found here
- Additional contributions are welcome!
- Please contact Cecilia (ceciliatosciri@uchicago.edu) and Joakim (joakim.olsson@cern.ch) if you are want to contribute