# Draft Survey Questionnaires for the DUNE Accelerator Integration Effort

—----------START OF SURVEY—----------

**This survey aims to gather information on algorithms, data models and memory management impacts the DUNE workflow . Your input will help improve accelerator integration strategies in the DUNE data reconstruction workflow.**

1. Please Specify the name of the workflow.
   ……

2. What Project do you work on the DUNE experiment?
   - Proto-DUNE
   - DUNE-FD
   - NDLAr
   - Other (Please specify)

3. What type of resource do you use for your workflow? *Check all that applies*
   - SuperComputer Facility (NERSC, ANL, FNAL, CERN resources)
   - Institutional Computing Resources (University clusters, local HPC)
   - Other (Example Cloud computing)

4. Accelerator types used in your workflow? *Check all that applies*
   - GPUs (Please specify the architecture)
   - N/A (Please Specify)

5. Percentage of the total compute time in the GPU? *Give your best estimation if such studies are not done yet*
   - 0%
   - 0-10%
   - 10-20%
   - 20-40%

- 40-80%
- >80%

6. What is the structure of the data-model used in GPU targeted tasks of your workflow (Select all that applies)
    - Flat arrays
    - Structure of Arrays
    - Arrays of Structures
    - Sparse Data representations
    - Custom format
    - Others (please specify)

7. What is the typical size of the input data product(s) that is offloaded in the GPU (per task)?
    - < 5 GB
    - 5-10  GB
    - > 10 GB

8. Is the Persistable data GPU friendly? (Data can be read and directly offloaded into GPUs without any transformation.)
    - Yes
    - No

9. What is/are the GPU targeted languages used in your workflow?
    - CUDA
    - Kokkos
    - SyCL
    - OpenMP
    - RAJA
    - HIP
    - Others (please specify)

10. What are the primary performance bottlenecks you experience? *(Select all that apply.)*
    - High Memory usage

- Computationally intensive tasks
- Poor parallel scaling of tasks
- Others (please specify).

11. Do you encounter memory leaks or excessive memory consumption in your current data processing framework?
    - Yes, frequently
    - Occasionally
    - No

12. Which of the following memory allocation strategies do you use? *(Select all that apply.)*
    - Pre-allocated fixed size device buffers to reduce runtime memory allocation overhead
    - Dynamic memory allocation inside Kernels
    - Dynamic memory allocation outside Kernels
    - Shared memory inside GPU kernels to reduce global memory access overhead
    - Unified Memory (Eg: cudaMallocManaged) for automatic access of memory between CPU and GPU
    - Others (please specify)

13. What are the challenges you face with your current data model? *(Select all that apply.)*
    - Data model (in ROOT, HDF5) is not GPU-friendly (for example AoS), needs to go through a transformation
    - GPU to CPU memory (and vice versa) data transfer bottlenecks
    - Poor scaling in different GPU environments (example works well in Nvidia, poor in AMD GPUs)
    - Other (please specify)

14. What tools do you use for profiling your GPU usage in the workflow (Example : Nvidia insight tools)
    - Please Specify

15. Do you do scaling tests before introducing new algorithms or data models in your workflow/software?
    - Yes
    - No

16. Does your software have a CI/CD pipeline to test the reproducibility of algorithms and data models?
    - Yes
    - No

17. Would a cookbook and/or test framework to validate data models and algorithms in GPUs within a standalone environment be helpful?
    - Yes
    - Maybe
    - No

18. Algorithm Identification: Which of the following algorithm types in your workflow do you believe are the best candidates for accelerator integration? (Select all that apply)
    - Data filtering and selection
    - Signal Processing
    - Simulation
    - Reconstruction
    - Pattern recognition
    - Other (Please specify)

19. Algorithm Bottlenecks: Among the following operations, which ones have you found to be performance bottlenecks in your accelerator integration?
    - Data I/O operations
    - Preprocessing or transformation algorithms
    - Post-processing tasks
    - Other (Please specify)

20. Do you have any examples or case studies where accelerating a particular algorithm significantly improved performance? Please share details in the space below:

21. Any additional comments or feedback? (*Please share any other issues not covered by the survey*)