Annual Report 2025
High Throughput Phenotyping and Data Storage Task Force (HTP-DS)

Current organizers

Claudia Kasper, Agroscope, Switzerland, claudia.kasper-voelkl@agroscope.admin.ch
Graham Plastow, University of Alberta, Canada, plastow@ualberta.ca
Cedric Gondro, Michigan State University, USA, gondroce@msu.edu

HTP-DS TF priorities

1. Identify and collate information on big data technologies and their uses in animal production
2. Identify community needs on High Performance Computing and develop resources for big data and artificial intelligence analytics
3. Create standards for reusability of data from sensors and imaging
4. Develop blockchain solutions for data integration across supply chains
5. Develop data storage methods for genomics and computational backbone for analyses (in collaboration with FAANGPredictions and FarmGTEx)
6. Establish a resource of genotypes, high-throughput phenotypes and environmental parameters for research
7. Develop a web server and accompanying portal to host data structures, programs, and resources

Status and plans

The task force has had limited coordinated activity in 2025. While progress has occurred on some priorities, particularly priorities four, five, and six, these advances have largely been driven by individual member efforts or by integration of results from external collaborative projects. As noted in last year's report, the task force continues to face challenges in securing dedicated funding and would benefit from closer alignment with synergistic FAANG task forces, particularly FAANGPrediction and MetaFAIR. In November this year, we held separate discussions with both groups to identify opportunities for better integration, and there was shared interest in closer coordination to align activities, reduce duplication, and increase momentum across overlapping areas of scope.

Given these constraints, it has become clear that the current set of task force priorities is too broad and ambitious for an unfunded and relatively unstructured small group to address effectively. If the task force is to make any form of measurable progress, its objectives need to be re-evaluated and consolidated around a single, well-defined vision that is better aligned with what FAANG (Functional Annotation of Animal Genomes) currently does,

namely the generation and structuring of functional genomic resources. The future plan is to leverage FAANG outputs into a more streamlined data science application area, using functional annotation and associated omics data as the backbone for advanced analytics and genomic prediction, rather than attempting to cover the full range of big data topics in animal production.

Building on these discussions, and in coordination with MetaFAIR (with FAANGPrediction intended as a core participant, pending further discussion), we plan to submit a USDA workshop grant proposal to support a structured, cross–task force planning process. The workshop will be designed to assemble a strong group of academics to develop the conceptual and organizational framework for a future SAS-scale project targeted for 2027, and produce a white paper articulating a shared vision of future pangenome applications, infrastructure requirements, and methodological needs for pangenome-anchored genomic prediction and multi-omics integration. At a high level, the anticipated SAS scope would center on pangenome-anchored genomic and functional annotation resources as core infrastructure, coupled with efficient data compression and storage strategies, AI-enabled and multi-omics-informed prediction methods, and interoperable data management consistent with FAIR principles. This reenvisioning should help develop a more coherent scientific direction, a refined and more intentional set of task force objectives, and a structured roadmap and timeline to support the subsequent SAS proposal development.