# Phenonaut; a tool for multiomic data integration, analysis, prediction, and inference

Steven Shave[1,2], John C Dawson[1], Richard Kasprowicz[2] and Neil O Carragher[1]

[1]College of Medicine and Veterinary Medicine, University of Edinburgh, The Institute of Genetics and Cancer, Cancer Research UK Edinburgh Centre, Crewe Road South, Edinburgh, EH4 2XR, UK,

[2]GlaxoSmithKline Medicines Research Centre, Gunnels Wood Rd, Stevenage SG1 2NY, UK.

Data integration workflows for multiparametric phenotypic data take many forms across academia and industry. Efforts with limited resources and smaller teams often encountered in academia can easily fall short of data integration best practices for processing and combining high content imaging, proteomics, metabolomics and other omics data. The additional difficulty in applying literature techniques with example code in disparate repositories, languages and compute environments means new techniques require significant investment before use. We present Phenonaut, a Python software package designed to address the data workflow needs of integration, migration, automation, and auditability in the application of literature and proprietary techniques for data source and structure agnostic workflow creation. The application of Phenonaut for omics analysis by non-computational specialists is exemplified by several use cases to demonstrate how the framework can be used for knowledge extraction from diverse biological datasets.