Biomedical research is increasingly driven by large-scale datasets: vast quantities of genome and EST sequences, epigenetic analyses, genetic polymorphisms and epidemiological data from diverse isolates, transcript and protein expression profiling results from multiple distinct platforms, phenotyping data of diverse provenance and reliability, automated and manually-curated annotation, information on protein structure, interactions, metabolic pathways, antigenicity, etc. Help!!! How can we most effectively collect, store, maintain, integrate, and mine these important datasets so as to define targets for further investigation? The Eukaryotic Pathogen Genome Database (http://EuPathDB.org) provides the protozoan parasite research community with convenient access to genomic-scale datasets, facilitating computational experiments that highlight and prioritize candidates for further study at the laboratory bench. Novel algorithms integrate diverse datasets to improve gene model identification and the recognition of targeting signals. Accurate identification of orthologs facilitates functional inference for ‘hypothetical proteins’. A new, portable, graphically-oriented user interface greatly simplifies the formulation and optimization of complex queries that can be shared with colleagues and stored for future review and refinement or modification. For example, users seeking novel vaccine targets may wish to prioritize (i) probable antigens that (ii) display evidence of evolutionary selection, (iii) are expressed at appropriate times (iv) in immunologically accessible locations, and (v) are conserved among pathogens but distinct from host species. Such queries support systems-level analysis of key biological problems, several of which will be illustrated through live demonstrations. This meeting will also provide ample opportunity for discussion of the philosophy, algorithms, and architecture underlying the EuPathDB.org, OrthoMCL.org and TDRtargets.org databases, if warranted.