Interest in network data has exploded in the last few years, with a host of applications such as the study of social networks, protein interaction networks, and the Internet. This has required the development of new generative models and new simulation methods. Yet it is usually infeasible to observe the full network, so sampling within the network is generally required.

Various network sampling schemes have thus been proposed. For example, respondent-driven sampling (RDS) is a link-tracing sampling design currently being used in public health studies all over the world. The standard RDS estimators rely on a long list of strong assumptions, such as that the degree of each individual in the sample is known exactly and that this individual is equally likely to recruit any of his or her neighbors in an underlying social network, and recent work of Goel and Salganik (and others) has shown that the widely-used interval estimates for these estimators can be wildly overoptimistic.

Previous work on sampling within a network and on modeling the network itself has mostly evolved separately and in parallel, but to obtain valid results for real networks it is crucial to understand more about how the model for the network meshes with how the sample within the network is obtained. Classical issues of model-based vs. design-based inference arise here, with new challenges from having also to account for the network structure. We will discuss these interactions and challenges, especially for the case of RDS, where several forms of uncertainty about the network sampling and the network model come together.