**Data Collection**: For three categories of forking: Undesirable and socially-related forking (U.F.), 2) Other socially-related forking (H.F.) 3) No forking at all (as the control group) (No.F.)

**Data Cleaning, Wrangling**: Forming longitudinal sociograms (directed graphs) using evenly-spaced snapshots of the run-up to the forking date for each project in all categories

**Statistical Modeling** using the developer-oriented statistical model.

Assume longitudinal evolution of network data is the result of many small atomic changes (ministeps) occurring between the consecutively observed snapshots of the network (graph)

Find the rate at which developers change one of their ties using the *rate function*.

Find the particular type of ministeps the developers make, using the *objective function and gratification function*. a) forming a new tie, b) breaking off an existing tie, c) maintaining a non-connection, d) maintaining a connection.

**Model Specification** using structural effects and behavior-related effects, for *Objective and Gratification* functions: Examples include *reciprocity effect, closure effects, three-cycles, density, betweenness effect, activity, similarity and assortativity effects*

*Model Simulation and Estimation: Estimate the model parameters and find a well-fitting model using Markov Chain Monte Carlo Estimation (MCMC).*

**Hypothesis Testing**: Now we have a well-fitting statistical model that captures the longitudinal evolution of sociograms of each project in all three forking categories, test the statistical significance of each model parameter using a single-parameter t-type test.
Compare two models and test the differences between two group, using a multi-parameter t-type test.
Compare categories of forking, using Multivariate Analysis of Variance Between Multiple Groups (MANOVA).