# How complex does your model need to be?

When we model a social process (like traffic or radicalization), we sooner or later come across questions like the following:

- How many choices do agents need to have?
- How many alternative theories/explanations should we test with this model?
- Which (if any) variables/parameters should be given real (i.e. empirically observed) values?
- When should we allow for some random noise (e.g. in agent's choice or in their environment)?
- How much of the parameter space should we explore with our simulations?

In your team project, some of these questions will arise as soon as you embark designing your model; some will emerge later, when you will start running the simulation experiments.
Notice how there isn't any generally valid answer to any of these questions: how we address them is, largely, an *arbitrary decision*. This leaflet is meant to make you aware of these decisions, so that you can make the best choices when designing your model.
When we talk about models in this leaflet, we generally refer to the ODD models [Grimm et.al.] which are a de facto standard in social simulation. ODD models define a number of important elements that you should describe in a model, but do not indicate the exact content or give a formal representation of it. They contain things like the purpose of the simulation, the agents, parameters that are needed, etc.

There are three factors usually playing a role in determining which arbitrary choice to make: the level of abstraction of our model, the granularity of its predictions, and some pragmatic considerations.

**Level of abstraction**
Models can vary between being very abstract and very realistic. Which kind you need depends on the *aim of your research*: let's take a model of criminal networks as an example. If the most important aim is to test alternative theories or to find out the predicted consequences of alternative network interventions, then the model can be more abstract than realistic. Agents' choices and attribute will be rather simplistic, and you will be more interested in trying out random permutations of the initial configuration of the model, than in modeling as accurately as possible an actually existing criminal network.
By contrast, you aim could be to develop possible network interventions to tackle specific (kinds of) criminal networks - for example, Dutch biker gangs. This aim leads to a more realistic model than before: empirical evidence will be of higher importance, and you will mainly focus on specific kinds of network topologies. In this example, you will focus on small, dense networks only (a sparse networks of 500 individuals is probably not representative of how biker gangs are organized).

**Granularity of predictions**
By granularity we typically refer to the *level of aggregation* of your model's outcome. We can illustrate this with the example of the electric cars model. We can think of different outcome variables for this model. For instance, an outcome variable could be the proportion of electric cars in circulation. This is an example of a variable measured at the neighborhood level, and thus at a high level of aggregation. Another variable could be the level of satisfaction/happiness of individual residents - this would be measured at the individual level, the lowest level of aggregation. There can be intermediate levels of aggregation, too (e.g. family/household level, street level, etc.).

In sum, what kinds of predictions you want your model to make, ultimately determines how much detail your model needs to have: if you want (or need) your model to make predictions on individual behavior (low level of aggregation), you will probably need to explicitly model individuals as agents.

**Pragmatic considerations**

When dealing with questions like the ones we listed above, it's important to keep in mind that we have limited resources. Adding new features to the model, adding more data, or exploring more alternative theories or parameter configurations takes time. Time is an important constraint to what we can do with our models, and thus may force us to make a pragmatic choice.

Grimm V., Berger U., DeAngelis D. L., Polhill G., Giske J.& Railsback S. F.. (2010) The ODD protocol: a review and first update. *Ecol. Model.* 221, 2760–2768.