ADEQUACY: WHAT MAKES A SIMULATION GOOD ENOUGH?

Samarth Swarup

Biocomplexity Institute & Initiative
University of Virginia
Charlottesville, VA, USA
swarup@virginia.edu

ABSTRACT
We examine several notions of what makes a simulation adequate, i.e., good enough for application. Social simulations, or artificial societies, are complex systems that integrate data from multiple sources, model intricate behaviors and processes, and generate large amounts of data. Adequacy, thus, is a nuanced concept that has to be applied at every stage of the development of the simulation, from initial conceptualization to final use. We go through this process step by step, and suggest some best practices for developing adequate models, drawn from experience with developing large-scale social simulations in multiple domains.

Keywords: social simulation, artificial societies, adequacy, synthetic populations

1 INTRODUCTION
A common (mis)conception is that the goal of modeling and simulation efforts should be to design the perfect simulation. This would presumably be a simulation which allows for perfect forecasting accuracy, so that we can feed in the current state of the world and determine what the future state of the world will be. We could then use the simulation as a black box and concern ourselves with just the inputs and outputs of the system. Even if this were possible, there are at least two fundamental problems with the idea.

The first is the measurement problem. Implicit in the idea of using a perfect simulation is the idea that we can provide perfect initial conditions; otherwise the simulation would not be able to provide accurate forecasts. However, this is impossible for any reasonably complex simulation. If we are constructing an artificial society that is a model of a real society—a city like Chicago, say—then it is obvious that we have no way of capturing a precise snapshot of the state of the city as a whole. Data collection takes time, and the city is changing even as the data are being collected. Further, for a complex distributed system like a city we do not have one mode of data collection; rather, data are collected by a multitude of administrative agencies, companies, official surveys, social media, physical sensors, and more. These are not coordinated efforts and result in data sets that are not designed to be integrated. Third, a city is not a closed system. There are myriad outside influences acting on a city that contribute to its operation and evolution. Providing perfect initial conditions would require accurate measurement of all these influences, potentially greatly expanding the scope of the simulation.

The second problem is the intervention discovery problem. A perfect simulation of a social system would necessarily be very complex, to provide requisite variety (Ashby 1958, Bar-Yam 2004). If the goal of
building a simulation is to provide some measure of control over the system, then we would like to be able to use the simulation to discover effective interventions. A rigorous approach to intervention discovery would require an optimization procedure that runs the simulation judiciously and searches through the space of possible interventions to find the best one. The more complex the simulation, however, the harder this optimization problem would become, to find the right interventions that can be used to steer the system in a preferred direction. A perfect simulation, then, would be nigh impossible to use in a practical sense.

While these two problems are stated in the context of “perfect” simulations, they do not rely on notions of computing with infinite precision or measuring the state of the world to the quantum level, etc. These are, instead, two very practical problems that arise even if we are trying to make “as-perfect-as-possible” simulations. However, the goal is not to argue for the futility of building artificial societies and simulations entirely. Rather, it is to step back from a pursuit of perfection to a notion of adequacy (Marathe et al. 2014).

If we try to do that, the first question that comes up is, adequate for what? Second, equally important, how is adequacy to be defined?

In what follows, we go through the process of the design and development of social simulations step by step and consider what adequacy means at each stage. We highlight a series of best practices, which are then summarized at the end.

2 THE GUIDING QUESTION(S)

If any simulation we construct is only going to be adequate for some specific purposes, this implies that we should have a clear understanding of the questions we are trying to answer before designing the simulation (Burton and Obel 1999). These questions will inform data collection, simulation design, and analysis of results. Conversely, limitations of data, design, and analysis will determine the caveats to the answers we are able to provide.

There are many reasons to do modeling and simulation, other than forecasting (Epstein 2008). For instance, we may wish to explain some observed phenomenon, understand the relative impact of different parameters or structures, quantify the uncertainty in outcomes, or to test a hypothesis about a specific real-world phenomenon.

For all these cases and more, questions can be asked at different levels of specificity. While it’s not always the case that the more specific the question is, the better, the question should be specific enough to be operationalizable. This means that it should be possible to implement a computational procedure that can answer the question. Stated as a best practice, a model should be designed to answer an operational question for which data are available. This can be illustrated through an example. Lum et al. (2014) developed a social simulation to evaluate if incarceration can be modeled as a contagious process. Though mass incarceration in the USA has been referred to as an epidemic (e.g., Gopnik 2012), a rigorous examination of this metaphor requires a detailed, data-driven model.

Lum et al. asked a specific question: Can the six-fold disparity in the incarceration rates of black and white Americans (Carson and Mulako-Wangota 2013) be explained by the much smaller disparity (≈20%; Bonczar et al. 2011) in black and white sentence lengths, by assuming that incarceration spreads as a contagious process? They were able to operationalize this question into a data-driven model of an evolving social network where all relevant parameters were obtained from public data, including data that could be turned into probabilities of contagion of incarceration by relationship-type (Dallaire 2007).

They were able to show that their model reproduced the observed six-fold disparity between black and white incarceration rates very closely, due to the epidemic threshold associated with an SIS contagion model. The difference in average sentence lengths situated the black and white populations on opposite sides of
the epidemic threshold, resulting in the “epidemic” of incarceration becoming endemic only in the black population. The model was validated by comparing distributions of recidivism rates, as predicted by the model, with actual data from multiple US states.

The important point to be made here is that this model was constructed to be adequate for answering a very specific question regarding the mechanism by which incarceration may grow to endemic levels in a population. It has limited applicability outside this context. For example, we may wish to evaluate potential interventions or policies for reducing the disparity in incarceration, in which case this model can only answer questions very closely related to the original question.

One such related question is, would ensuring that future sentence lengths for all convicts are chosen from the “white” sentence length distribution be sufficient to eliminate the disparity in incarceration rates between blacks and whites? It turns out, according to the simulation, that this would reduce but not completely eliminate the disparity in incarceration rates (Hawdon et al. 2017). This is because clustering in the social network creates local “reservoirs,” which means that individuals keep passing the contagion to each other and the incarceration rate stabilizes at a higher value than it would in a population that has always had sentence lengths assigned from the “white” sentence length distribution. If we wish to know how to reduce the disparity further, the model can only suggest that this clustering should be eliminated from the social network, which is not a realistic suggestion since clustering is an innate feature of social organization.

There are other possibilities for eliminating the disparity, of course, such as retraining/reintegration programs, increased social support for the families of those incarcerated, and more, but this model is not adequate for evaluating those types of policies or interventions. In fact, the model suggests a kind of data collection that is needed, viz., data about the effects of those kinds of programs on the probability of contagion of incarceration. Were those data available, it might be possible to construct a more complex and detailed model that would be adequate for evaluating the relative effects of those interventions and policies. Without that, it would be an improper use of the model to try to generate policies for reducing racial disparities in incarceration.

The context of a model refers to the larger social system in which the phenomenon of interest is embedded. When we build a model to answer a specific, operationalizable question, we necessarily leave out distantly related or tangential aspects of the larger social system. A model which is adequate in its limited, specific setting may not be applicable to answering questions about other aspects of the larger social system. Stated as a best practice, a model should not be used outside the context for which it is adequate.

3 WHAT TO LEAVE OUT

Given an appropriate context or guiding question, we have to make choices about abstraction and representation. Abstraction is the process of choosing which aspects of the real-world phenomenon of interest are to be modeled, and which can be safely ignored. This is a difficult question in general and the answers depend on a careful examination of the real-world phenomenon and a deep understanding of the domain (often through consultation with domain experts).

Representation refers to the data structures and rules that implement the abstracted phenomenon in a computer simulation. The choice of representation has important implications for multiple aspects of the process: it affects computability and scalability of the simulation, the input data needed to construct the representation, and the kinds of actionable information that can be generated from the simulation. For instance, if we choose a representation where every agent has to make a decision at every time step, even if most agents actually act only intermittently, that may result in poor scaling performance. Second, if data are not available directly about the chosen representation, then a process is needed to construct the representation from the
data. Third, the representation determines the kinds of interventions that can be evaluated, as we saw in the previous section also. We illustrate these issues through examples from computational epidemiology.

If we want to study the spread of an infectious disease like the flu at a population scale, we need to represent the interaction process through which flu spreads from person to person in a population. This interaction depends on physical proximity and duration of contact, about which it is difficult, if not impossible, to get data directly. This is simply because most people don’t know all the people they are in physical proximity with during the course of a typical day, since such interactions occur in workplaces, neighborhoods, stores, public transport, etc., with a population much larger than a person’s circle of acquaintances.

For this problem, it is now common to use an artificial society model known as a synthetic population (Adiga et al. 2015, Eubank et al. 2004). A synthetic population represents every individual in a region, along with their typical daily activity patterns as well as the locations where these activities are carried out. From this, an approximation to the social contact network can be constructed. The resulting networks are structurally more naturalistic than common stylized network models (Karra et al. 2018) and exhibit differences due to the differences in contingent realities of different cities and regions (Xia et al. 2013, Xia et al. 2014). Synthetic populations are used for many applications in addition to epidemiology; see Müller and Axhausen (2010) for an overview.

With complex models like synthetic populations, there are a number of decisions regarding abstraction and representation that are made, which affect the adequacy of the simulations for different questions. For example, the choice to leave out some segment of the population, such as slums (Adiga et al. 2018) or tourists (Parikh et al. 2013) can have a significant impact on the assessment of epidemic severity or the effectiveness of mitigation efforts. Similarly, a choice to represent the interactions as a static graph can lead to a significantly faster simulation (Bisset et al. 2009) as compared to a representation based on activity patterns and mobility (Barrett et al. 2008), but at the cost of reduced flexibility to represent individual decision-making and adaptation.

In each case, choices of abstraction and representation need to be guided by the motivating questions about the phenomenon of interest, and these in turn determine the design of an adequate model and simulation. Similarly, the impact of the abstraction and representation on the outcomes guides the determination of adequacy also. For example, the most common use-case is to generate a ranking of potential interventions, whether they be pre-determined or discovered through the simulation itself, not accuracy of prediction. In other words, policy-makers and planners need to know what is the best thing to do, not necessarily exactly what the outcome will be. Thus a simulation might be considered adequate if it generates the right ordering of interventions, even if it doesn’t generate precise point predictions (e.g., Halloran et al. 2008). Stated as a best practice, careful consideration should be given to choices of abstraction and representation in order to construct an adequate simulation of the phenomenon of interest.

4 SYNTHETIC DATA DESIGN DECISIONS

Once a context, abstraction, and representation have been chosen, the artificial society has to be constructed, often by combining data from multiple sources. For example, a synthetic population is generated in a series of steps, where each step involves integrating data from an additional source (Adiga et al. 2015). The process is briefly described below.

1. **Demographics**: A population of synthetic agents and households is generated by combining data of two types from the American Community Survey: summary data, which are marginal distributions of relevant demographic variables in the region, and the corresponding Public Use Microdata Sample (PUMS), which is a 5% sample of anonymized survey records.
2. Activity patterns: Activity sequence data from the National Household Travel Survey are integrated with the population, to assign a typical daily or weekly activity sequence to each agent.

3. Activity locations: Locations are assigned for each activity for each agent, taking into account the capacities and types of the locations, generally derived from business location surveys and the National Center for Education Statistics.

A synthetic social contact network is generated from the synthetic population by choosing a model for interaction between agents who are at the same location for an overlapping time period (Karra et al. 2018). At each step of the process outlined above, choices have to be made that determine the adequacy of the resulting synthetic population in terms of the questions and domains to which the model can be applied (Marathe et al. 2014). Some important choices are as follows.

Granularity: What constitutes an individual agent in the artificial society? Depending on the domain, it might be, e.g., a person (Eubank et al. 2004), a household (Zhang et al. 2016), or a firm (Axtell 2016).

Resolution: The level of detail in the representation. For example, is age represented in years, in years and months, in years, months, and days, etc.; is space measured in kilometers, meters, centimeters, etc.; what is the duration that corresponds to a time step of the simulation, and more.

Fidelity: The number of properties that are represented. For example, are people represented by their age, their age and gender, their age, gender, and income, etc.

Accuracy: At what level of aggregation is the representation accurate and to what degree? For example, the representation may be statistically equivalent to the real population at the level of Census block groups, but not at a higher resolution than that. This can be the consequence of the accuracy of the input data, e.g., the US Census does not release data publicly at the block level for privacy reasons, or it may be a design choice influenced by the level of aggregation at which outcomes are desired.

Different combinations of choices lead to synthetic data sets that are adequate for different problem domains. For instance, a number of household-level details, such as the size of the house, the heating fuel, whether the house has a pool, etc., may be necessary to include if the goal is predict the probability that the house will adopt solar panel technology (Zhang et al. 2016). However details of the daily activity sequences of the household members may not be relevant. On the other hand, if the goal is to model household energy demand, then daily activity sequences provide useful information about active energy usage (Thorve et al. 2018). Thus, two different contexts in the energy domain can require very different choices of granularity, resolution, and fidelity to construct adequate simulations. Stated as a best practice, aspects of synthetic data, such as granularity, resolution, fidelity, and accuracy should be chosen to match the requirements of the context.

5 DATA AND INFORMATION QUALITY

In addition to the above properties, there are several dimensions of data and information quality that are applicable to synthetic information. Quality broadly refers to the usability of the data and information, and is a broad area of research (e.g., Batini et al. 2009, Stvilia et al. 2007). There are many metrics of data and information quality, but some new ones are needed for synthetic information (Gupta 2014). This is because when we synthesize information by integrating data from multiple sources, we can produce information which cannot be directly measured for the purposes of validation. An example is the synthetic social contact network, as discussed earlier. Further, since a synthetic population is constructed in multiple steps, it is advisable to have measures of data and information quality at every stage of the process. Some relevant dimensions are defined below.
Completeness: This refers to the absence of missing values or missing entities. For an artificial society, this means, e.g., that every person in the region of interest should be represented, and that all the agents should have values for all attributes that are included in the representation.

Consistency: This refers to both consistency of variables describing an individual as well as consistency across individuals. The issue is exacerbated when we are combining multiple data sources into a synthetic representation. For example, travel distances, which are generated by the process of assigning locations for activities should be consistent with travel times, which are part of the activity sequences that are assigned to synthetic individuals.

Validity: This refers to ensuring that the values assigned to attributes should be in appropriate ranges. For instance, if the synthetic population ages over the course of a simulation, then we should have an appropriate “death” mechanism to ensure that ages stay bounded. Validity also refers to constraints on states or relationships in a synthetic population. For example, there are often laws against leaving children under a certain age alone at home, so the activity sequences that are assigned to members of a household should be such that this constraint is not violated.

It should be noted that this term in the information quality literature is drawn from the domain of database research, and has a different meaning from ‘validation’ of simulations, which refers to determining the correctness of simulations by comparing outcomes against real-world data.

Reliability: Since a synthetic population is the outcome of a stochastic process, it is important to bound the variability in the process so that the resulting populations are not too different. Reliability can be evaluated by comparing synthetic populations generated by the multiple runs of the process. This is, in a sense, the converse of our earlier point about determining adequacy of abstraction and representation through examination of the variability in the outcomes of the simulation. The point here is that trust in a simulation is easier to establish if we feed it realistic data produced in a reliable manner. If the process of producing synthetic populations generates highly variable populations, it casts doubt on the realism and generalizability of the simulation results.

Timeliness/Currency: This refers to keeping the synthetic population up to date as the underlying data sources get updated. Different data sets get updated on different schedules, which can also introduce inconsistencies. Thus currency might have to be traded off with consistency.

The adequacy of a model and simulation depends on all these dimensions and others. A lack of quality along one or more of these dimensions doesn’t automatically render the model inadequate. That depends, as before, on the magnitude of the effect on outcomes. For example, it might be admissible to have some missing values or to have some invalid activity sequences (such as ones that leave some children alone at home for short periods), if they do not significantly affect the epicurves (the number of new infections over time) in an epidemiological study. Establishing this, however, requires a careful sensitivity analysis. Stated as a best practice, a careful sensitivity analysis should be done to determine appropriate levels of tolerance along multiple dimensions of data and information quality to assess adequacy. For large and complex simulations, it may not be possible to use Monte Carlo methods for sensitivity analysis. In such cases, emulation methods (response surface methods) can be used (e.g., Fadikar et al. 2017), though much work remains to be done in this area.

6  UPDATING

So far we have mostly discussed static aspects of the design of artificial societies. We now turn to dynamic aspects. The timescale of the simulation is a crucial aspect of the design. An adequate timescale is one which allows the smallest unit of effective change to be modeled. For example, a detailed traffic simulation may need to represent driving decisions at the sub-second level, whereas an infectious disease epidemic
simulation may only include interactions that are above a threshold that gives a reasonable probability of infection (say, five minutes). The choice of timescale also depends on the temporal resolution of the input data. For example, the activity times and durations available through an activity survey might be rounded to the nearest half-hour.

Another, related, design choice has to do with the update sequence. In a simulation, at a given time step, the order in which agents are updated can have a profound impact on the phase space of the simulation (Barrett et al. 2000). This means that the set of equilibria towards which the simulation evolves can change based on the update order. In order to eliminate this variability, simulations often adopt a parallel (synchronous) update mechanism, where each agent can only access the state of the simulation at the previous time step when updating to the current time step (effective parallelism), or where agents are updated in random order. Determining the full impact of update order on a large simulation is very difficult and is an active area of research (e.g., Laubenbacher et al. 2009, Rosenkrantz et al. 2018).

Updating refers to any change in state of the agents, including learning. The update mechanisms implemented in the simulation lead to particular trajectories and attractors in the state space of the simulation. The choice of the mechanism, therefore, directly affects the dynamical phenomena observed in the simulation, and determines the realism of the simulation. This dependence is brittle if changes to the update mechanism cause large changes in the outcomes. Therefore the robustness of the outcomes of the simulation should be tested by reasonable variations to the update rules, to show that the simulation adequately reflects the causal processes thought to be operating in the real world.

In some simple cases, it is possible to establish the robustness of an outcome of interest to multiple update schemes as a way of assessing the adequacy of the simulation procedure. A classic example is the Schelling (1971) model of neighborhood segregation. In his model, households exist on a grid, and the neighborhood of a household consists of the eight neighboring cells on the grid. A household can be of one of two races, and has a preference for living in a grid cell where at least a threshold number of its neighbors are of the same race. Schelling shows, through experimentation, that for multiple different update schemes, neighborhoods get segregated over time, even for fairly low thresholds.

This exceedingly simple, stylized model sparked an enormous amount of research in computational sociology because Schelling was able to use it to make an important point: that though there are many overt and intentional reasons neighborhood racial segregation occurs, there may also be a systemic, unintentional mechanism behind this phenomenon that is driven by small individual choices that aggregate over time to lead to an undesirable global situation. Thus this is a classic example of a model being used to explain rather than predict, despite the fact that it meets almost none of the criteria for adequacy discussed in previous sections. Its success lies in convincing the reader that it is capturing an essential causal mechanism at operation in the world that leads to the observed phenomenon.

Stated as a best practice, an adequate simulation must convincingly capture at least some essential causal mechanisms driving the phenomenon of interest in its update mechanism. It is not necessary to represent all the causal mechanisms at play, as Schelling demonstrates, if the goal is to isolate one or more causal mechanisms and investigate only their effects.

7 SIMULATION ANALYTICS

Once we have a complete, operational model of an artificial society, where all the previous questions have been resolved, we can ask, can we now answer the questions for which the simulation has been designed? More importantly, can we explain the answers generated by the simulation? This is essential to building trust in the simulation, and thereby establishing its adequacy. However, the larger and more complex a model gets, the more it becomes opaque to inspection. Large-scale simulations have the limitation that we can only
run them a few times in a reasonable amount of time, which means we cannot explore their parameter spaces completely, nor can we run large-scale statistical experiments, since these can even overwhelm reasonable HPC systems. Sense-making in this regime is a challenge, and new methods are needed. Each simulation run can generate a huge amount of data, so that even summarizing the results of a single run can pose an interesting technical challenge (Parikh et al. 2016).

Interactions in a simulation result in emergent effects, such as distributed agency and distributed (socio-logical) structure. Simulation analytics refers to the methods being developed to answer questions using complex large-scale simulations (Swarup et al. 2018). For example, an end-user of a simulation may ultimately care about what the right thing to do is, in a particular scenario. Ideally, the user should be able to give a partial query, such as “what should people do one hour after a disaster occurs?” The simulation system should be able to intelligently complete this partial query and give an answer like “one hour after the disaster, people who are in areas A and B should do X, while people in other areas should do Y.” This kind of behavior ranking capability (Parikh et al. 2017) contributes to the adequacy of a simulation system that goes beyond properties of the data or the structure of the simulation.

The larger point here is that adequacy of a simulation is also dependent on the uses to which the simulation can be put, which are determined by additional analytic capabilities that must be built to make the simulation usable by operational end-users. In section 3, we discussed the “what” of simulation design, i.e., what the question is that we are trying to answer. The point here is about the “how” of simulation design, i.e., how the simulation enables the user to find the answer. A user who is not technically sophisticated in simulation science can still make effective use of a simulation that provides mechanisms for investigating its workings in a user-friendly way. This is a lot like software development, where querying capabilities, interface design, and other affordances (interoperability, robustness, etc.) are as crucial to success as correct and reliable operation.

Stated as a best practice, an adequate simulation should provide the means for effective use by domain experts and end-users.

8 ETHICAL AND MORAL CONSIDERATIONS

Lastly, and perhaps most importantly, we should be aware of the human values that are implicitly or explicitly embedded in the design and use of any model of a social system (Friedman et al. 2006). Values, in the form of biases, can be present in data collection systems. While survey design is a mature field of study, today’s big data systems gather data from a variety of unstructured and uncontrolled data sources as well. Unless we are very careful, this can result in privileging some segments of society over others in the artificial societies we create from these data sets.

Similarly, in the processes and causal mechanisms we choose to model, and in the experiments we choose to design with our models, biases can be present. For example, if we choose to evaluate interventions in a disaster that consist of sending emergency messages on cell phones, as compared to sending emergency personnel knocking on doors, we might find that the former has better outcomes because it reaches a larger number of people more quickly. However, it will also systematically miss people who are either too poor to have cell phones or prefer not to use them, e.g., the elderly. Value-sensitive design, in this case, would require us to think about whether we value the greatest good for the greatest number (utilitarianism) or a certain minimum standard of care for everyone (let’s call it dignity).

In the process of moving from data to policy, there is a gap. While models can provide evidence for policy, making the jump to the actual policy to be implemented requires making decisions about how to use the models. It is in this process that our values are most starkly manifest. The best model and simulation will prove to be inadequate if it is not used in the right way.
Stated as a best practice, **an adequate simulation is one that is used for ethical and moral good.**

9 CONCLUSION

Here we pull together the best practices listed in the previous sections:

- A model should be designed to answer an operational question for which data are available.
- A model should not be used outside the context for which it is adequate.
- Careful consideration should be given to choices of abstraction and representation in order to construct an adequate simulation of the phenomenon of interest.
- Aspects of synthetic data, such as granularity, resolution, fidelity, and accuracy should be chosen to match the requirements of the context.
- A careful sensitivity analysis should be done to determine appropriate levels of tolerance along multiple dimensions of data and information quality to assess adequacy.
- An adequate simulation must convincingly capture at least some essential causal mechanisms driving the phenomenon of interest in its update mechanism.
- An adequate simulation should provide the means for effective use by domain experts and end-users.
- An adequate simulation is one that is used for ethical and moral good.

While this article develops these notions of adequacy in order from the conceptualization of the problem onward to the use of the simulation, this is by no means a linear process in practice. There is much feedback between stages. For example, it may turn out that when we are trying to implement essential causal mechanisms, we realize that the representation needs to change. Or it may turn out that effective use by domain experts requires a change in the scope of the simulation. Designing and implementing an adequate simulation, therefore, is an iterative process that, ideally, includes the end-users as well.

These best practices, and the concept of adequacy in general, are meant to drive the development of practical, usable simulations. These are not meant to be the only, or definitive, set of best practices. The concept of adequacy, likewise, is expected to grow and evolve as the community discovers new problems as well as new solutions and methods in simulation science. There are other recent overviews of the praxis of agent-based modeling and simulation, which focus on different aspects. For example, O’Sullivan et al. (2012) discuss when the complexity of developing an agent-based simulation outweighs the benefits, Grimm et al. (2005) present a particular methodology, called pattern-oriented modeling, to the development of agent-based models, Abdou et al. (2012) present a set of steps to follow while developing an agent-based model, and Macal (2016) gives a general overview of the state of the art in the field. The present article, with its focus on adequacy, is complementary to these overviews, adding another facet to the ongoing development of the philosophy of agent-based modeling and simulation.

There are a number of important research directions to be explored. Broadly, more effort has been devoted to the earlier stages of the simulation design process than the later stages. Thus, while there is much work in simulation formalisms, architectures, and scalability, there is relatively little in understanding causality in simulations, simulation analytics, and the ethics of simulations. With the ongoing emergence of data science and fairness, accountability, and transparency (FAT*) as a field of study, there is also a need for a deeper understanding of how these issues permeate into models and simulations and consequently into plans and policy.
ACKNOWLEDGMENTS

I thank Wesley Wildman and Saikou Diallo for the invitation to participate in the round table on Best Practices in Building Artificial Societies at SpringSim 2019. I thank Philippe Giabbanelli and Kathleen Carley for very helpful comments on an earlier draft of this paper. I also thank the members of the Network Systems Science and Advanced Computing Division of the Biocomplexity Institute at the University of Virginia for numerous helpful discussions. This work was done in part while the author was at Virginia Tech. This work is supported in part by Air Force Research Laboratory Contract FA8650-18-C-7826, Next Century Corporation (federal flow through from DARPA) Agreement NC082817-VT-GORDIAN (from Contract No. FA8750-17-C-0155), DOE Cooperative Agreement DE-EE0007660, DTRA CNIMS Contract HDTRA1-17-0118, NIH Grant 1R01GM109718, NSF Grant CCMI-1745207, NSF 16-610 Smart and Connected Communities (S&CC) Grant CNS-1737492, NSF IBSS Grant SMA-1520359, and NASA grant 80NSSC18K1594.

The views, opinions and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

REFERENCES


Swarup


Karra, K., S. Swarup, and J. Graham. 2018, Dec. “An Empirical Assessment of the Complexity and Realism of Synthetic Social Contact Networks”. In GTA³ 2.0: The 2nd workshop on Graph Techniques for Adversarial Activity Analytics. Seattle, WA, USA.


AUTHOR BIOGRAPHY

SAMARTH SWARUP is a Research Associate Professor in the Biocomplexity Institute & Initiative at the University of Virginia. He holds a PhD in Computer Science from the University of Illinois at Urbana-Champaign. His research interests include large-scale social simulation using synthetic populations, human behavior modeling, and simulation analytics. His email address is swarup@virginia.edu.