



Birkin, M., Allan, R., Beckhofer, S., Buchan, I., Finch, J., Goble, C., Hudson-Smith, A., Lambert, P., Procter, R., de Roure, D. and Sinnott, R.O. (2010) *The elements of a computational infrastructure for social simulation*. Royal Society of London Philosophical Transactions A: Mathematical, Physical and Engineering Sciences, 368 (1925). pp. 3797-3812. ISSN 1364-503X

<http://eprints.gla.ac.uk/7452/>

Deposited on: 31 August 2010

# The Elements of a Computational Infrastructure for Social Simulation

Mark Birkin<sup>1</sup>, Rob Allan<sup>2</sup>, Sean Beckhofer<sup>3</sup>, Iain Buchan<sup>4</sup>, June Finch<sup>5</sup>, Carole Goble<sup>3</sup>, Andy Hudson-Smith<sup>6</sup>, Paul Lambert<sup>7</sup>, Rob Procter<sup>5</sup>, David de Roure<sup>8</sup>, Richard Sinnott<sup>9</sup>

[1] School of Geography, University of Leeds; [m.h.birkin@leeds.ac.uk](mailto:m.h.birkin@leeds.ac.uk) (corresponding author) [2] STFC, Daresbury [3] School of Computer Science, University of Manchester [4] School of Medicine, University of Manchester [5] School of Social Sciences, University of Manchester [6] Centre for Applied Spatial Analysis, UCL [7] Applied Social Science, University of Stirling [8] Electronics and Computer Science, University of Southampton [9] NeSC, University of Glasgow

## 1 Introduction

Social simulation is a relatively new and rapidly expanding field with applications across the academic research community. Both micro-simulation and agent-based modelling have been widely adopted within economics, sociology and geography. Simulation models have also provoked high levels of interest in healthcare research, anthropology and political science. Policy interest is substantial at local and national government levels in a range of forecasting scenarios, including transport, housing, education and healthcare, and within the private sector.

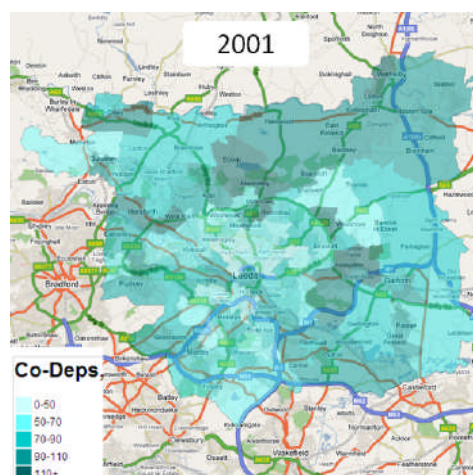
The authors of this paper are leading efforts to construct a National e-Infrastructure for Social Simulation (NeISS) in the UK. This paper will present a review of the objectives and plans of the NeISS infrastructure, and outline current technical approaches and challenges. This £1.5M project, which is funded under the JISC Information Environment programme, is both important and timely for a number of reasons: there is a **growing demand** for social simulation ‘services’ among the UK academic community, and in the public and private sectors; while the UK National Centre for e-Social Science (NCeSS) offers a increasingly substantial **critical mass** of expertise. Some key NCeSS projects which are partners in NeISS are reviewed in Section 2 of this abstract. The dissemination of tools and services is backed through a network of **international collaborations** involving both European and Asian Grid Infrastructures; and solid **foundations** have been laid for this ambitious programme, which we refer to in Section 3 of the abstract. There is demonstrable user engagement which has been stimulated by **ongoing innovations** in e-social simulation, while buy-in from the research community is assured through the prospect of **return on existing investments** in both computational and data infrastructures.

## 2 Objectives of the NeISS Infrastructure

The MoSeS NCeSS Node at the University of Leeds (Models and Simulation in e-Social science) is the precursor to NeISS [1]. MoSeS is concerned with the projection of simulated city and regional populations into the medium-term future, from 2006 until 2031. There are no existing projections which support planning the delivery of services like housing, healthcare and transportation at the intra-urban scale.

Consider the example in Figure 1. A category of co-dependent households has been identified which are of critical importance for the delivery of social care (co-dependent households are elderly couples living together in which one household member is heavily dependent upon the other). Through detailed simulation modelling the current and future distribution of such households has been determined.

The model has three components: a population reconstruction method; a dynamic forecasting model; and a scenario-based (what if?) impact analysis tool with social and behavioural modelling functionality. MoSeS social care simulations combine data from the ONS mid-year population estimates and sub-national population projections, census data,



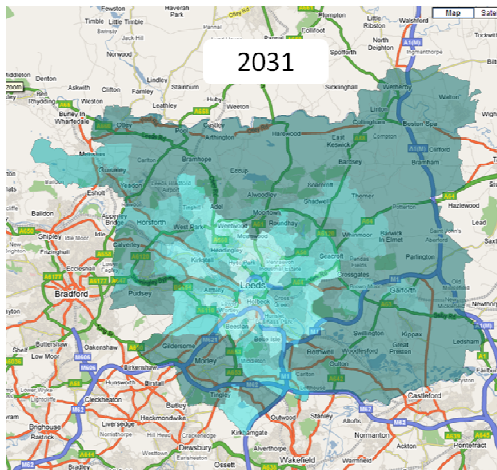


Figure 1. Simulation of co-dependency in Leeds

*Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland. 2001 Census, Output Area Boundaries. Crown copyright 2003.*

longitudinal micro-data from the British Household Panel Survey and third party information from Leeds City Council Social Services and the Leeds Teaching Hospitals Trust.

NeISS also builds on contributions from three other NCESS nodes: GeoVUE, DAMES, and Obesity e-Lab. GeoVUE is concerned with the provision of extended visualisation capabilities for analysis and validation of simulation model outputs using Web2.0 mapping tools, and building on existing work which encourages crowd-sourcing of data [2]. The DAMES Node focuses upon the support and documentation of activities in manipulating and preparing data for analysis. Its input to the NeISS infrastructure will involve embedding tools for documentation, standardisation and enhancement of data management activities within the simulation resource. This will provide environments to allow researchers not simply to supply their data to the simulation service, but also to perform reviews and manipulations of that data in a manner which is consistent with social science data standards. Obesity e-Lab is a multi-disciplinary team operating across Health, Social and Computer Sciences. This activity is heavily application focused, enabling social and biomedical researchers to share data, information and analytical tools for obesity research.

The contributions of the various NCESS Nodes described above are combined flexibly to produce social simulation workflows which include primary datasets, simulation code, tools for analysis and

visualisation of scenario outputs and archives for the deposition of results from simulation experiments or scenarios (see Figure 2). In the next section, the computational means for the composition and enactment of such workflows will be discussed.

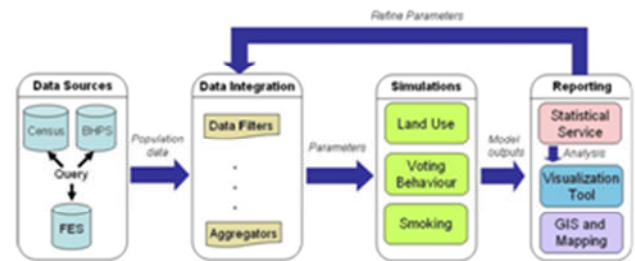


Figure 2. Example social simulation workflows

### 3 Computational Requirements

Ne-ISS consists of four interacting layers (see Figure 3). The first is a service layer comprising the fundamental components on which simulation depends, and which have been described in Section 2 above. The second is a composition layer in which individual services are composed into workflows and then published as coherent entities. The integration of Taverna functionality into a Portal [3] is planned to provide JSR 168/286 compliant components that provide access to enactment, management, monitoring and creation of workflows. This provides the means by which social simulation services can be combined into research-oriented 'activities' as well as providing a means for the introduction of new functionality.

The publishing strand draws on myExperiment [4], a VRE for the social curation and sharing of scientific Research Objects, especially workflows and in silico experiments, facilitating their integration into a portal. An important focus is the incorporation of social networking functionality to provide extended archiving and documentation capability for social simulations. Research Objects, aggregations of resources (data sets, analysis methods, workflows, results, people) that tell a particular story about an investigation, experiment or process and capture key information about the lifecycle of the investigation (for example, provenance information about analyses), facilitating re-use of results and repeatability of experiments. In combination, layers one and two provide the basis for the integrated support of the 'social simulation research lifecycle' in which intelligence is generated from data, published, shared, analysed, archived and perpetually recycled.

Above the composition layer sits an architecture layer. This provides tools and methods needed to provide portal access to simulation services and workflows, and to combine these into domain-specific exemplars. NCeSS has an existing Sakai community portal. Alternative technologies are also under consideration. The portal can be accessed through Internet2 Shibboleth technologies and UK access management federation. Tools will be provided to improve overall security and personalising of content based upon user privileges. Sakai is a VO management system, not just a portal and role-based access is part of its framework. It has fine-grained permissions for each tool based on a user's role within a realm and core services provide this information to Sakai tools. The issue is how to get this information to and from external services, e.g., using SAML. NeISS will be built around a Service Oriented Architecture into which core services and external services are designed to be pluggable, using frameworks and standards developed under the e-Framework for Education and Research and OMII-UK. Core services including data and computation will be supported through existing e-infrastructures, including resources hosted by the project partners at STFC Daresbury, White Rose Grid, and North West Grid.

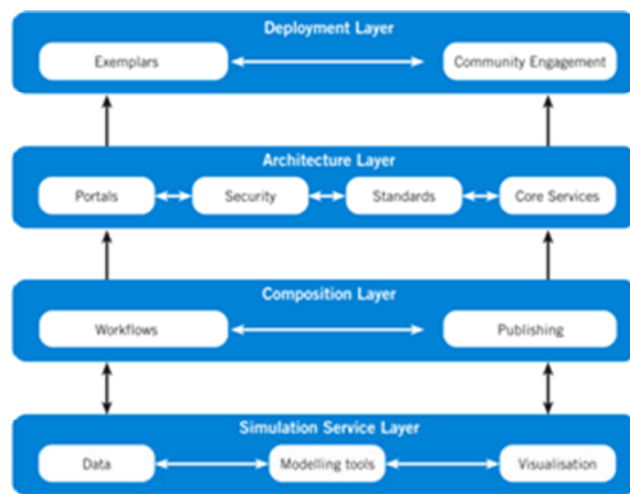


Figure 3. Component layers in NeISS

In the deployment layer, Manchester's e-Research Centre is coordinating a range of activities to facilitate delivery of applications and increase take-up within both the social science community and applied users. The social simulation providers introduced in Section 2 have an established foothold in housing, transport, health and social care, education and demographic planning. Engagement with inexperienced users accessing services which are integrated within domain-specific service users will provide a clear demonstration of the impact of this

work. However the success of the project will also be conditioned by its ability to engage with experienced users accessing capabilities at the level of an individual service or workflow; and expert users with the capability to introduce their own services, and build new workflows through combinations of new or existing services. In addition to their practical value, exemplars will be used as a means to focus component services into more meaningful workflows, research objects and portlets which address particular social science research challenges. The exemplars will draw on dependencies between all of the other e-infrastructure components and therefore serve as a focus for the integration of development activities between all groups.

#### 4 Conclusions

NeISS is a unique project aiming to build simulation services for social scientists and policy-makers. The project is characterised by an interest in the deployment of simulation tools for data integration, modelling, analysis and simulation. These tools are manipulated in an Information Environment (portal) which facilitates other activities in the research life-cycle, such as the composition and enactment of workflows, publication and sharing of analytical results.

In our presentation to the All Hands conference we will illustrate the service components of NeISS in relation to a specific social care application. The requirements of users, data integration, analytical functionality and infrastructure components will all be considered in relation to this specific context.

#### Acknowledgements

NeISS is funded under the JISC Information Environments Programme.

DAMES, GENeSIS and Obesity e-Lab are all funded as nodes in the ESRC National Centre for e-Social Science.

#### References

- [1] Townend P, Xu J, Birkin M, Turner A, Wu B (2009). Modelling and Simulation for e-Social Science, *Philosophical Transactions of the Royal Society A*, 367, 2781-2792.
- [2] <http://www.maptube.org>
- [3] <http://www.taverna.org.uk>
- [4] <http://www.myexperiment.org>