Towards a Prototype Policy Laboratory for Simulating Innovation Networks

Michel Schilperoord, Petra Ahrweiler
Innovation Research Unit, University College Dublin, Ireland
e-mail: michel.schilperoord@ucd.ie

Abstract

This paper presents an approach for designing and building a computational laboratory for research and innovation policy simulation, centred around the SKIN model. The aim of the paper is to bring together empirical and computational research for policy use. The SKIN model will be embedded in a workflow and an interfacing infrastructure that supports rich user interaction with the lab's simulation database.

Introduction

In the past few years, the SKIN model has been applied to various studies simulating the impacts of different policy choices on the structure, formation and effectiveness of research and innovation networks. This has answered an explicit demand of policymakers and evaluation experts in this area, who asked for policy impact simulation "to make evaluations relevant to policy options for intervention in RTD and innovation. Evaluations must relate observed parameters and impacts to the characteristics of the intervention. It must be possible to deduce what could and should be changed in the intervention to improve impacts. Accordingly, much more use should be made of ex-ante network analysis to simulate the impacts of intervention policy changes" (European Commission Workshop Report "Using Network Analysis to Assess Systemic Impacts of Research" [1]).

For example, working from datasets on EU-funded research networks in the area of information and communication technologies [2, 3], the INFSO-SKIN application, specifically-developed for the Directorate General Information Society and Media of the European Commission (DG INFSO), helped to understand and manage the relationship between research funding and the goals of EU policy. The agents of the INFSO-SKIN application are research institutions such as universities, large diversified firms or small and medium-sized enterprises (SMEs). The model simulated real-world activity in which the Calls of the Commission specify the minimum number of partners, composition of partners, and the length of the project; the deadline for submission; a range of capabilities, a sufficient number of which must appear in an eligible proposal; and the number of projects that will be funded. The rules of interaction and decision implemented in the model corresponded to Framework Programme (FP) rules; to increase the usefulness for policy designers, the names of the rules corresponded closely to Framework Programme terminology. For the Calls 1-6 that had occurred in FP7, the model used empirical information on the number of participants and the number of funded projects, together with data on project size (as measured by participant numbers), duration and average funding. Analysis of this information produced data on the functioning of, and relationships within, actual collaborative networks within the context of the Framework Programme. Using this data in the model provided a good match with the empirical data from EU-funded ICT networks in FP7: the model accurately reflected what actually happened and could be used as a test bed for potential policy choices.

Altering elements of the model that equate to policy interventions such as the amount of funding, the size of consortia, or encouraging specific sections of the research community, enabled the use of SKIN as a tool for modelling and evaluating the results of specific interactions between policies, funding strategies and agents. Because changing parameters within the model is analogous to applying different policy options in the real world, the model could be used to examine the likely real-world effects of different policy options before they were implemented.

Another example is the application IPSE-SKIN [4], which is currently developed in the context of national innovation policies to answer the needs of the Irish economy to breed an innovation ecosystem with optimally structured university-industry-government networks. Informed by large new datasets coming from empirical research on the Irish national innovation system, IPSE-SKIN is used to investigate the actors, designs, processes and policies of Irish innovation networks for developing and testing policy strategies for Irish innovation networks, and work on options for anticipating and analysing new developments to help the recovery of the economy. The high degree of empirical information in the model, together with the strong applied-orientation of experiments to examine how research networks change and adapt in response to different policy stimuli demonstrate the utility of this approach for policy makers.

In both examples, communication of results to the sponsors, clients and users of the studies was strongly supported by a graphical user interface that allowed policy makers to perform their own experiments without the requirement to directly change the code. This offered a kind of “lab experience”, which was greatly appreciated from the user perspective.

However, the experiences from these applications and the confrontation with stakeholder and user demands, which became obvious during the studies, now ask for a more systematic approach to use the SKIN model for research and innovation policy impact simulation. We need to provide an integrated infrastructure of services and interfaces that can be offered to lab users. Our final aim is to provide lab facilities on the Internet by means of easy-to-use web apps with data visualisations designed for interaction with detailed data extracted from the lab’s simulation database. How can this be realised, and what are the concrete steps to achieve this?

The conceptual set-up and the technical realisation must tackle both, internal issues of standardization, quality control, and effectiveness, as well as external features such as visualization and communication with and for lab users. The internal processes refer to the workflow of the lab and to the question how the SKIN model is exactly embedded in it.

In this paper, we present the computational designs developed so far for building a research and innovation policy laboratory in silico in order to inform policymakers about optimal network structures for research and innovation.

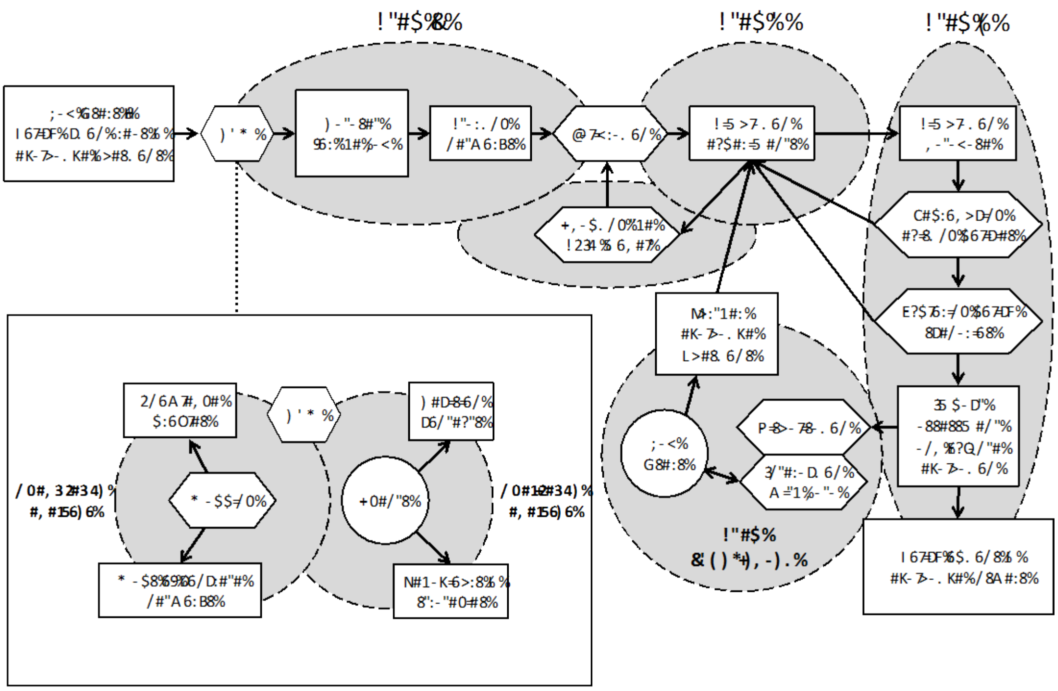
Realistic Simulation

The stakeholder-driven purpose of the lab requires to implement and test innovation policy or innovation management scenarios in a realistic and evidence-based way. The users of the lab are mostly not interested in results for scientific reasons, or wonder about the more general features and mechanisms of research and innovation networks. Their questions concern the SKIN application of their concrete network under investigation, which should therefore be informed by empirical data and be implemented with much care to detail.

This often asks for handling large databases mapping the knowledge profiles and agent characteristics of research and innovation networks. Furthermore, adapting the SKIN model to various case studies of research and innovation networks requires qualitative analyses of decision contexts, agent behaviours and strategies and their implementation into the model. Only if the model features empirically observed processes leading to the empirically observed structures mimicking the causal mechanisms of real life, model-suggested interventions into these processes, in terms of empirical policy changes, can change empirical network structures in an expected way. Therefore, we are not only looking for SKIN reproducing empirical network structures but also reproducing the ways these are generated. The model needs to work with empirical information on both.

To find realistic starting configurations for simulating research and innovation networks under investigation, we need a close link between the model and this empirical information involving a high fraction of calibrated distributions concerning features of the agents and specific context conditions such as existing policies at work.

Figure 1. Lab workflow and data-to-model (D2M) strategy



The lab prototype builds on a three-step data-to-model strategy:

1. Use empirical data on research and innovation networks as data input to calibrate the SKIN model adapted to the study context (“starting networks”);
2. Process this data in simulation experiments with different policy scenarios relevant for the lab users (mostly network monitoring, ex-post evaluation, impact assessment, ex-ante evaluation);
3. Analyse and interpret the simulated network data and provide results for lab users.

Calibration

If we have empirical information as outlined available, we can apply our framework for transforming data points to model input deciding which categories of the databases can be used to set up the model.

While qualitative data will mostly inform the agent behaviours and the rules of the system, quantitative data can be used for calibrating the structure of the research and innovation network under investigation. Our calibration procedure tries to match the structure of the ‘real’, empirically observed, research and innovation network with the structure of the ‘artificial’ network generated by the simulation model, before we even start with any simulations, i.e. we just try “to reproduce the database” with the model. The database is connected to the model using MySQL, which means that data can be read directly from within the model with standard SQL queries (the interface has a switch called “empirical case” to do this).

Figure 2. Calibration: empirical case mode vs. simulation mode

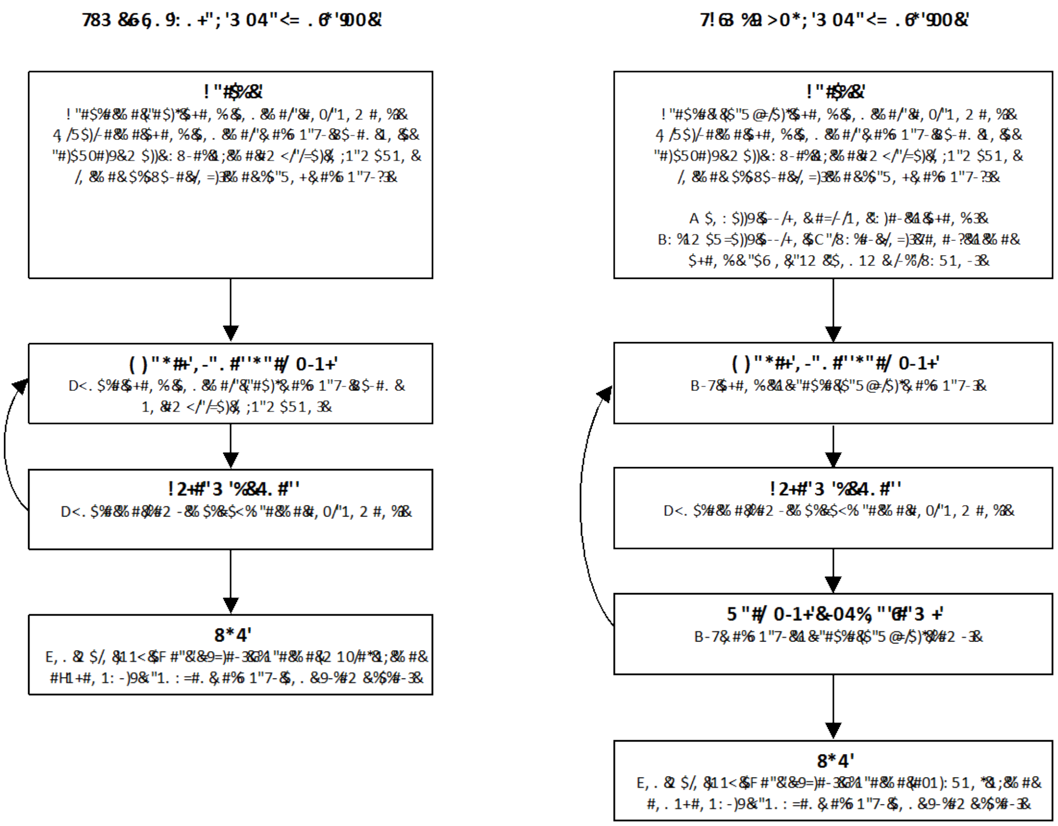


Figure 2 shows how the main loop differs when running the model in empirical case mode. In this mode the agents are ‘real’ in the sense that they directly point to the network participants stored in the database (e.g. company X or university Y). In set-up, a relatively small subset of the information on starting networks in the database is used to create and initialise ‘real’ networks between these agents. These networks are ‘real’ in the sense that they directly point to the networks stored in the database (e.g. partnership Z). After set-up, each cycle of running the model amounts to updating the agents and the networks based on empirical information in the database. This stepping through the cycles is similar to playing back the “movie” of the agents creating these networks. The beauty of this approach is that it allows us to use many of the procedures that are also used when running the model in simulation mode, e.g. the procedures for system updates (each cycle) and storing simulation outputs (at the end of the last cycle). This makes calibration much easier.

The calibration procedure is:

1. Switch the platform into ‘empirical case’ mode.
2. Set-up the case. This starts the initialisation of the model using the empirical database. Agents and environment are created based on the information in the database.
3. Start the SKIN application. The number of iterations corresponds to the time covered in the database (if time is recorded).

4. Each iteration, the platform updates its internal representation of the network of the 'real' agents, which are all directly linked to information on actual organisations in the database. The updating process uses the empirical information retrieved from the database. This stepping process (not simulation) recreates the evolution of the ‘real’ research and innovation network and, taking advantage of the platform’s user interface and analytical features, we can observe how this network structure has evolved.
5. Switch the platform into ‘simulation’ mode.
6. Set-up the case. This starts the initialisation of the model using input parameters extracted from the empirical database. Agents and environment are automatically created based on these parameters. (In ‘simulation’ mode, some items might not be created at initialisation, since they will be outputs of the simulation)
7. Start the SKIN application.
8. Each iteration, the platform updates its internal representation of the network of the agents, based on the endogenised procedures (which might be calibrated by qualitative data, see above), i.e. created by the agents and outcome of the policy strategies.

The ‘real’ and ‘artificial’ networks, automatically stored for each iteration in series of (GEXF) data files, can be compared using the complete set of network analysis tools, either using the algorithms that are built into the platform, or from software libraries like Gephi, NetworkX or other. If we are confident that we have reproduced the major stylised facts of the research and innovation networks under investigation that have been extracted from empirical information, the model can be used as a laboratory to change the design of instruments and to analyse divergent developments.

Workflow

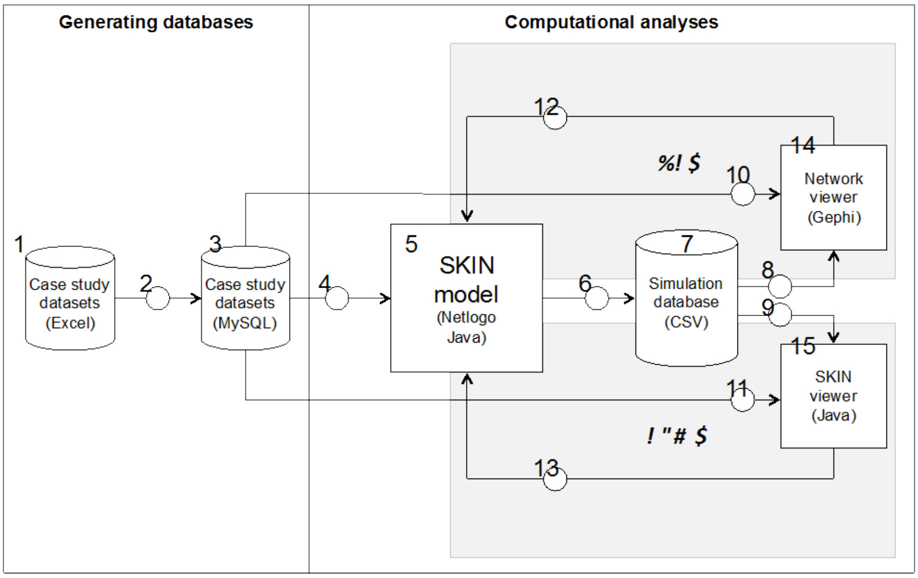
Workflow, meaning the often tedious interlinking of resources, is an important topic when it concerns empirical calibration and validation. Assessing the performance of models against observed data requires a constant review of the entire protocol for techniques used in constructing and interrogating (i.e. experimenting with) these models [5, 6]. The ABM community is to respond with standard techniques: workflows and protocols that are tailored to commonly used resources and techniques for constructing ABMs and integrate these with widely used ABM tools such as NetLogo and RePast. While standard techniques can be helpful, reality is that agent-based modellers will often need to adapt their workflows and protocols to very particular problems concerning their resources and techniques. So they also need to be creative. But how can one match and reconcile both needs: robustness and creativity?

Figure 3 displays an open adaptable software platform with a key role for the SKIN model surrounded by middleware – the software that glues things together. This platform was developed during the studies mentioned above and already applied and tested within these applications. It will be further elaborated and extended within current and future research. To develop our calibration and validation methodologies, more ways of interlinking component workflows with middleware need to be envisioned and tested.

The currently existing middleware is designed to *(i)* interlink the lab components (the empirical datasets, models, an integrated simulation database, analytical tools and visualisations) in a flexible manner, *(ii)* deal with data-related problems, *(iii)* facilitate a general workflow that is explorative and data-driven, and *(iv)* exploit the ABM’s strengths by covering an as wide as possible range of use cases, including unforeseen use cases.

For calibrating the SKIN model with empirical data on research and innovation networks, we developed software for bridging between databases and the SKIN model. This involved formatting the network data in such a way that it can be read into the SKIN model in order to structure the initial network of agents at the start of the simulation.

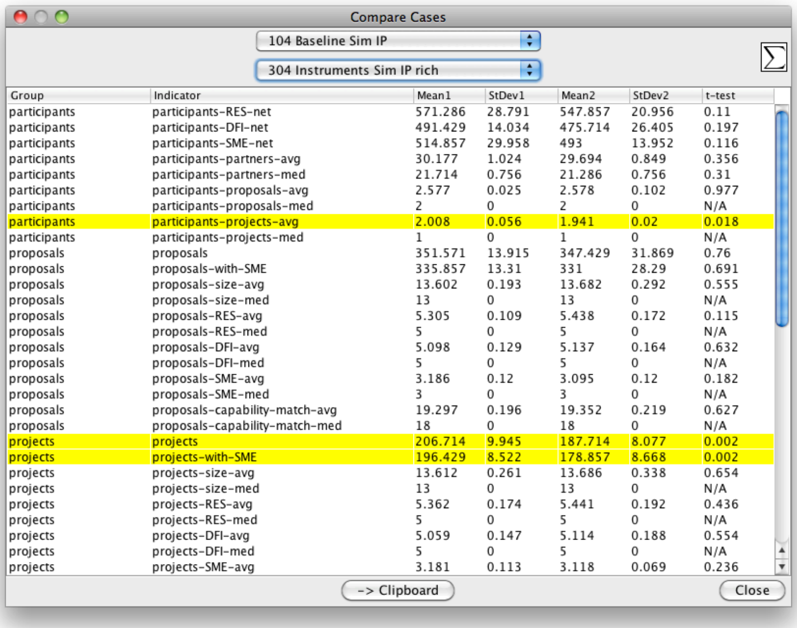
Figure 3. General workflow



Middleware links are displayed as 2, 4, 6, 8, 9, 10, 11, 12 and 13

The visualisation and evaluation of the results stored in the simulation database is much helped by the SKIN Viewer tool, which was programmed in Java for the INFSO-SKIN application. It provides several functions to visualise and compare the large volumes of data generated in experiments with the INFSO-SKIN model, including a t-test for statistically significance of findings. It can be easily adapted to other SKIN applications.

Figure 4. Screen shot of the SKIN Viewer tool – summary table with t-test results



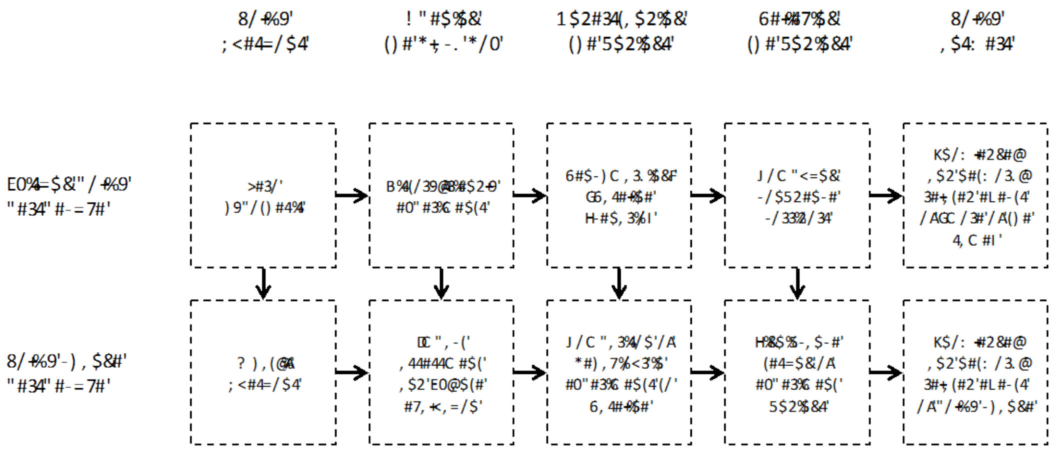
Group	Indicator	Mean1	StDev1	Mean2	StDev2	t-test
participants	participants-RES-net	571.286	28.791	547.857	20.956	0.11
participants	participants-DFI-net	491.429	14.034	475.714	26.405	0.197
participants	participants-SME-net	514.857	29.958	493	13.952	0.116
participants	participants-partners-avg	30.177	1.024	29.694	0.849	0.356
participants	participants-partners-med	21.714	0.756	21.286	0.756	0.31
participants	participants-proposals-avg	2.577	0.025	2.578	0.102	0.977
participants	participants-proposals-med	2	0	2	0	N/A
participants	participants-projects-avg	2.008	0.056	1.941	0.02	0.018
participants	participants-projects-med	1	0	1	0	N/A
proposals	proposals	351.571	13.915	347.429	31.869	0.76
proposals	proposals-with-SME	335.857	13.31	331	28.29	0.691
proposals	proposals-size-avg	13.602	0.193	13.682	0.292	0.555
proposals	proposals-size-med	13	0	13	0	N/A
proposals	proposals-RES-avg	5.305	0.109	5.438	0.172	0.115
proposals	proposals-RES-med	5	0	5	0	N/A
proposals	proposals-DFI-avg	5.098	0.129	5.137	0.164	0.632
proposals	proposals-DFI-med	5	0	5	0	N/A
proposals	proposals-SME-avg	3.186	0.12	3.095	0.12	0.182
proposals	proposals-SME-med	3	0	3	0	N/A
proposals	proposals-capability-match-avg	19.297	0.196	19.352	0.219	0.627
proposals	proposals-capability-match-med	18	0	18	0	N/A
projects	projects	206.714	9.945	187.714	8.077	0.002
projects	projects-with-SME	196.429	8.522	178.857	8.668	0.002
projects	projects-size-avg	13.612	0.261	13.686	0.338	0.654
projects	projects-size-med	13	0	13	0	N/A
projects	projects-RES-avg	5.362	0.174	5.441	0.192	0.436
projects	projects-RES-med	5	0	5	0	N/A
projects	projects-DFI-avg	5.059	0.147	5.114	0.188	0.554
projects	projects-DFI-med	5	0	5	0	N/A
projects	projects-SME-avg	3.181	0.113	3.118	0.069	0.236

The second issue the conceptual set-up and the technical realisation of the lab prototype must be able to deal with lab-external features such as visualisation and communication with and for lab users.

Interaction with the Lab: a User's Perspective

For a lab prototype on research and innovation policy we have to envisage that we will have to deal with simulation data from case studies that differ importantly in focus, scope and scale, yet have a common conceptual reference: policy questions concerning research and innovation networks.

Figure 5. The lab functions from a user's perspective



From a user's perspective there are three issues, where the first and the third are yet outside the scope of themes we already discussed, but which are central to the lab user if not the only issues any user is interested in:

(i) the policy-relevant questions the user wants the simulation to answer

(ii) the credibility of the findings supported by the evidence base of the realistic simulations calibrated by empirical data (often provided by the users themselves)

(iii) understanding the answers the simulation findings provide

Evaluative Questions

Depending on whether the workflow is applied to a tender study with a clear client demand behind it, or to a more open research project sponsored by a general funding agency, the questions the simulation needs to answer will be more or less pre-defined from the onset of a project.

However, we learnt that even in very client-driven studies, enough time should be dedicated to identifying and discussing the exact set of questions the stakeholders of the work want to see addressed. We found that the best way to do this is applying an iterative process of communication between study team and clients, where stakeholders learn about the scope and applicability of the methods, and where researchers get acquainted with the problems policy makers have to solve and with the kind of decisions, for which sound background information is needed. This iterative process will result in an agreed set of questions for the simulation, which will very often decisively differ from the set proposed at the start of the study. For the INFSO-SKIN study, a so-called "Steering Committee" was assigned to us consisting of policy makers and evaluation experts of DG INFSO. For the IPSE-SKIN application, we have employed a Stakeholder Board of high-profile Irish innovation policy representatives.

The strength of our type of methodology lies in the opportunity to ask what-if questions (ex-ante evaluation), an option, which is normally not easily available in the policy-making world. To find the right set of what-if questions, the benchmark question to go from is the “zero hypothesis”: what if there are zero changes? This just follows the extension of the time horizon for the simulated research and innovation networks without changing the rules (the so-called Baseline Scenario). To do this, just “more of the same” is made available to the agents and the networks. Answering the “zero hypothesis” question is important for two reasons: First, the original line of development and network evolution is extended and the sustainability of the network architectures is tested. Second, we can use this scenario as a kind of benchmark and compare its outcomes with the results of further experiments, as naturally no empirical data about the future are available for comparisons.

If the networks created by real life and those created by the agent-based model qualitatively correspond closely, the simulation experiments can be characterized as history-friendly experiments, which reproduce the empirical data and cover the decisive mechanisms and resulting dynamics of the real networks: ‘History-friendly models are formal models which aim to capture - in stylized form - qualitative theories about mechanisms and factors affecting industry evolution, technological advance and institutional change put forth by empirical research in industrial organization, in business organization and strategy, and in the histories of industries. They present empirical evidence and suggest powerful explanations. Usually these “histories” (…) are so rich and complex that only a simulation model can capture (at least in part) the substance, above all when verbal explanations imply non-linear dynamics’ [8].

This qualifies the model to be applied for in silico experiments, which take these outcomes as a starting point and test the potential future developments in the sense of the ex-ante evaluation outlined above. INFO-SKIN, for example, carried out a number of experiments to simulate policy options in Horizon 2020 and their consequences for the future ICT research landscape in Europe. The results were compared with the baseline scenario of the unmodified FP7 ICT Calls 1-6. Differences between the behaviour in the experiments and that in the baseline could be attributed to the effects of the policy change.

What types of questions are possible? As SKIN is about knowledge dynamics and inter-organisational network dynamics, evaluative questions can address effects of policy changes on two levels – knowledge and agents/networks.

The knowledge-related indicators [cf. 9] demonstrate the additional advantage of a simulation to record the development of an “un-observable”. Empirically, we cannot see and only with difficulty can we measure knowledge generation, knowledge exchange, and learning. In the simulation, the effects of knowledge exchange processes are measured as knowledge flows between the agents in the networks. With SKIN, we can see what policy changes do to the amount of knowledge in the system and the characteristics of the knowledge flows.

The network-related indicators show many statistical properties of networks that are standard applied in network analysis, thus allowing the same type of inquiry. Simulation offers the advantage of seeing effects of policy changes over many (new) instantiations of these networks, and therefore gives us information on their dynamic characteristics, e.g. are networks adaptive to changing conditions, robust, sustainable? Which agents are they attracting? Are they productive? (not only in terms of knowledge production but also, for example, in terms of creating niches within the larger network structure for accelerated growth). For the latter type of questions, the simulation's main loop (Figure 2) offers a facility – the "networks produce items" step – that can be adapted to particular applications.

Evaluative questions can address both, the knowledge and the network level. For example, the agreed set of evaluative questions for the INFSO-SKIN application only contained one question for the knowledge level (the first one) and various questions for the agents/networks level. These were:

- What if there are changes to the thematic areas, which are funded?
- What happens if there are changes in the funding instruments (small actions, large actions)?
- What if there are interventions concerning the scope/outreach of funding?
- What happens if there are changes in the funding level (increase, decrease)?
- What if there are interventions concerning the participation of certain actors in the network (e.g. SMEs)?

For the IPSE-SKIN application, questions also concern structure and operation of networks, their appropriate scale and scope (present technology/industry sectors and new fields) and how the networks are adapted and embedded in context of national, regional and local agendas. Evaluation concentrates on finding the optimal network structures for: *(i)* supporting enterprises and promoting competitiveness, *(ii)* strengthening the whole innovation cycle (entrepreneurship, innovation, commercialisation), *(iii)* stimulating the green sector (energy security, protecting the environment) and *(iv)* improving the efficiency and quality of public services and regulation (Building Ireland's Smart Economy [10]).

Helping users to understand and interpret the findings

Our user's perspective, geared towards asking evaluative questions, puts heavy demands on the lab's interfaces system, not only for being able to adapt the lab to specific questions, but also for allowing users to understand and trust the findings. We already pointed out middleware as a flexible way of adapting the lab and setting up a workflow that is generally robust. Our experience is that interaction by users with data from network analysis (NA) and simulations (ABM) can also help creating a more solid basis for understanding and trust, adding to the traditional technical style of reporting findings.

The good news is that sophisticated ways of interfacing and interacting with NA/simulation data are increasingly feasible. The new possibilities are reflected and foreshadowed by the numerous extensions created by NA/ABM communities, as illustrated by growing number of Netlogo extensions and Gephi plug-ins. They can build on technical developments in providing data-driven apps on the Internet with strong visualisations. Possibilities are illustrated by MIT's Observatory of Economic Complexity [11].

To optimally benefit from these possibilities, the lab provides four interfaces:

- the users interface (UI) for remote access to the lab,
- the web applications interface (API) for bringing to life the lab's data,
- the modelling processes interface (MDI), including the middleware, for setting up and controlling computational processes, and
- the databases interface (DBI) for storage of large streams of data feeding in and out of the lab.

The UI facilitates remote access to the lab's web pages using desktop or mobile devices with a standard web browser. It connects to the API for quickly serving up web pages containing apps that enable visualisation of, and interaction with, the lab's data. Expert users of the lab use the MDI for setting up and controlling the lab's computational processes, whereas empirical research teams have their own interface for binding their data inputs to the lab. The DBI also offers tools for inspecting and managing the lab's results, which are automatically stored in a single, integrated simulation database.

The computational laboratory in silico can inform policy and management on optimal network structures for innovation performance adapted to contextual conditions. Our simulations will serve as a laboratory to experiment with social life in a way that we cannot do empirically due to methodological reasons [cf. 12]. Using this tool, we can understand innovation dynamics in complex social systems and find their potential for design, intervention and control.

References

1. European Commission Workshop Report "Using Network Analysis to Assess Systemic Impacts of Research". March 2009: 18
2. Ahrweiler, P., Schilperoord, M., Gilbert, N. and Pyka, A. (forthcoming). "Research policy laboratories. Agent-based modelling for impact assessment and ex-ante evaluation of complex policy instruments"
3. Ahrweiler, P., Schilperoord, M., Gilbert, N. and Pyka, A. (2012). Final report – "Using network analysis to monitor and track effects resulting from changes in policy intervention and instruments". SMART 2010-0025, DG INFSO, January 2012, Brussels
4. Ahrweiler, P. (2012). "Conceptual framework of computational policy research: a guiding document and reference for the IPSE consortium". Deliverable D1.1, Project report, <http://www.ipse.ie>
5. Windrum, P., Fagiolo, G., and Moneta, A. (2007). "Empirical validation of agent-based models: alternatives and prospects". *Journal of Artificial Societies and Social Simulation* 10(2)8, <http://jasss.soc.surrey.ac.uk/10/2/8.html>
6. Lorscheid, I., Heine, B.-O. and Meyer, M. (2012). "Opening the 'black box' of simulations: increased transparency and effective communication through the systematic design of experiments". *Comput Math Organ Theory*, March 2012, Volume 18, Issue 1, pp 22-62
7. Ahrweiler, P. (Ed.) (2010). *Innovation in Complex Social Systems*. London: Routledge
8. Malerba, F., Nelson, R., Orsenigo, L. and Winter, S. (1999). "History-friendly models of industry evolution: the computer industry". *Industrial and Corporate Change* 8, 3-40
9. Ahrweiler, P., Pyka, A. and Gilbert, N. (2011). "A new model for university-industry links in knowledge-based economies". *Journal of Product Innovation Management* 28, 218-235
10. Department of the Taoiseach (2008) Building Ireland's smart economy: a framework for sustainable economic renewal (Dublin, 2008)
11. Simoes, A. J. G. and Hidalgo, C. A. (2011). The Economic Complexity Observatory: an analytical tool for understanding the dynamics of economic development. Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence, <http://atlas.media.mit.edu>
12. Ahrweiler, P. and Gilbert, N. (2005). "Caffè nero: the evaluation of social simulation". *Journal of Artificial Societies and Social Simulation* 8(4)14, <http://jasss.soc.surrey.ac.uk/8/4/14.html>

Acknowledgement

We gratefully acknowledge the financial support from the Irish Government's Programme for Research in Third Level Institutions (PRTLI 5) project grant entitled 'Innovation Policy Simulation for the Smart Economy (IPSE)'.