Theory Project Proposal

**Project Title:** **ABM-based open source data collaboration and constraint model**

**Student Name: Zixiang Ma**

**Student ID: 2056647**

**Supervisor Name: Christopher Baber**

**Project Category/Topic:**

● Theory

**Project Aim:**

With the advancement of observation technology and information processing technology, massive spatial data has been acquired today. At the same time, with the development of open source technology and data storage technology, data can be stored online, and can be shared and exchanged globally, which also provides the possibility of collaboration between different people or companies. But unlimited means hidden dangers. The privacy data or copyright data of the data source may be maliciously stolen, and the security of the data source cannot be guaranteed. So, the real question is how can we share as much data as possible to allow collaboration while retaining control over who can access that data?

The project aims to build a model to study how unrestricted sharing of all forms of data can support collaboration, and on the other hand, the impact of such sharing on data source protection. This protection may involve data privacy, copyright and security, and relevant conclusions can be obtained through model experiments.

**Related work:**

Before starting work on the project, you need to understand and learn some models and knowledge related to the subject of this project. This project involves the dissemination and sharing of information, collaboration between different individuals or companies, and the need to introduce protection and restriction measures, so I downloaded the relevant models and conducted research. At the same time, because the project will use an agent-based model, I consulted some information about the agent-based model. The agent-based model is a computing model used to simulate the actions and interactions of autonomous agents (independent individuals or common groups, such as companies, countries), and evaluate the role of the agent in the overall system through image display. This is very consistent with the characteristics of this project.

**Project Objectives/Deliverables:**

1. Learn the theory of open-source data sharing
2. Learn to use NetLogo modeling tools
3. Research and learn related models (e.g., there are lots of information diffusion and 'gossip spread' models)
4. Successfully build a basic model
5. Add some restrictions or protections to the model
6. Add variables and run simulation experiments
7. Obtain reliable or credible experimental results
8. Verify that the protections or restrictions added by the model can successfully prevent criminals from acquiring all data or break protection
9. Complete the final project report

**Methodology:**

This project uses NetLogo programming modeling, I have downloaded and tried NetLogo. I think NetLogo is very convenient to use, the interface interaction is clear and concise, and it is very suitable for people like me who try to model for the first time. In addition, NetLogo also has its own official communication community and model resource library. Beginners can try many simple models and modify the models according to their own ideas. At the same time, they can also ask questions in the forum, and other users can solve them. your problem. After the model is established, add independent variables and dependent variables and conduct simulation experiments to obtain the results of the model evaluation. For example, the independent variable can be set to the level of data protection, the amount of data shared, and the dependent variable can be set to how quickly information spreads through a network with no constraints and what effects the introduction of constraints will have (either in terms of limiting spread of some information or in terms of the 'success' of the collaboration).

**Project plan:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Task/Milestone** | **Jun** | | **Jul** | | **Aug** | | **Sept** |
| **Log1-2** | **Log3-4** | **Log5-6** | **Log7-8** | **Log9-10** | **Log11-12** | **Log13** |
| **Literature review** |  |  |  |  |  |  |  |
| **Determine the project and proposal** |  |  |  |  |  |  |  |
| **Learn NetLogo and related models** |  |  |  |  |  |  |  |
| **Create the first generation model** |  |  |  |  |  |  |  |
| **Run and modify the model** |  |  |  |  |  |  |  |
| **Second generation model** |  |  |  |  |  |  |  |
| **Modify variable values and record model experiment results** |  |  |  |  |  |  |  |
| **Thesis writing** |  |  |  |  |  |  |  |

**Risks and contingency plan:**

Because it is the first time to work on modeling, it may take more time to study modeling software. At the same time, because I have not learned the basic knowledge related to data security, when building a model, I may not consider it comprehensively and need the guidance of my supervisor.

**Hardware/Software Resources**

This time I used the NetLogo modeling tool. NetLogo is the most popular ABM toolkit used in the socio-ecological modeling community. NetLogo is very friendly to beginners and it is very convenient to build a model. There are many model resources for reference and learning, which can help solve problems.

**Data**

1. Salamon, Tomas (2011). Design of Agent-Based Models : Developing Computer Simulations for a Better Understanding of Social Processes. Bruckner Publishing. ISBN 978-80-904661-1-1.
2. Bonabeau, Eric, Agent-based modeling: methods and techniques for simulating human systems. Proc. National Academy of Sciences 99(3): 7280-7287, 2002.
3. Axelrod, Robert, The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration, Princeton: Princeton University Press, 1997, ISBN 978-0-691-01567-5
4. http://ccl.northwestern.edu/NetLogo/models/index.cgi