Overview

**Introduction**

This chapter is the first part of this report and gives a brief overview of the whole report. It mainly includes the project problem to be solved, the background knowledge needed to understand the background of the project, the motivation for choosing the project, a clear and detailed description of the project goals and an overview of the solution structure.

The purpose of the work described in this report is to establish a model that simulates data sharing and collaboration, and to study the impact of some of the model's restrictions on data sharing and security. This method of controlling variables helps to get a perfect data sharing and collaboration model.

The second chapter of the thesis is the literature review part, which introduces the previous work in the academic literature related to this project and provides proof for the project, proving that the problem is a practical and important issue. The third chapter is the model realization part. This part introduces in detail the definition of the model structure and goals, the variables contained in the model, the functions implemented by the model, and subsequent improvements to the model. Readers can understand how the model does not even run the model by themselves. use. The fourth chapter is the model testing part, which introduces in detail the adjustment and selection of model variables, the results of model operation and the sample results obtained. The fifth chapter is the conclusions and shortcomings. This part introduces the results of the model operation in the form of graphs, and compares, analyzes and infers the results, so as to obtain convincing conclusions and proves the results before the model is established. Hypothesis. In addition, it also explained the current shortcomings of the project and the work to be done in the future.

1.1 Background

In the network age, the storage and dissemination of data has become a breeze. Data sharing is the foundation of the development of the digital economy and the digital transformation of social governance. The world is connected through huge network data, and through sharing data for all-round and diversified collaboration. For example, the European Union attaches great importance to network data sharing, and specifically proposes the "European Data Strategy" guidelines for this purpose. Among them, it is proposed to strengthen the top-level design of cross-departmental data use, improve the level of public sector data sharing, and use legal measures to promote corporate data sharing. In addition, there are research institutions such as health and medical institutions, document management institutions, etc., which have established the basic principles of international health and medical data sharing, scientific research data open sharing guidelines and other data sharing and collaboration guidelines. People all over the world can use data to cooperate or conduct research together without meeting, and the era of big data has also been deepened in different fields. In addition, this year's COVID-19 epidemic also proved the importance of data sharing and collaboration. After the outbreak of the epidemic, China notified the World Health Organization of the epidemic situation and provided the virus sample and structure test report, so that other countries can take precautions and prepare for the epidemic early. After the domestic epidemic was brought under control, China submitted a hundred-page epidemic report and response measures to the World Health Organization, such as blocking cities and streets and isolating suspected patients. Facts have proved that after some European countries (the United Kingdom, Serbia, etc.) have adopted corresponding measures, the epidemic has indeed been effectively controlled, and this is due to the data and experience shared by China. This also proves the importance and necessity of data sharing and collaboration in today's globalization.

But it is worth noting that no matter what field or type of data, it should be shared and collaborated under the premise of reasonableness, legality, and compliance. With the rapid progress of information technology, the leakage and transactions of personal information data are causing great troubles to people's lives. The importance of personal information protection has become more prominent, and so is data. However, due to imperfect privacy protection key technologies, imperfect privacy protection laws and regulations, insufficient privacy protection awareness, and insufficient industry self-discipline, in recent years, Internet technology-based crimes that violate data security and personal privacy have shown explosive growth. For this reason, the most important issue for organizations or institutions in various countries or regions is information and data security, and they have taken actions one after another. For example, the European Union, although it is difficult to compare Europe's competitiveness with China and the United States in the Internet industry, its role and role in the Internet field should not be underestimated and ignored. With the entry into force of the “General Data Protection Regulation” (GDPR) in May 2018 and the huge US$5 billion fine imposed on Google in July, Europe has been involved in data governance, super network platform antitrust, personal information and privacy protection and other current network governance issues. A series of measures taken in key areas shocked the world. As the Internet Governance Forum (IGF) has been held in Europe for four consecutive years, together with the World Summit on the Information Society in Geneva (WSIS), the European Internet Governance Dialogue (EuroDIG), and the ICANN CEO position held by Europeans, they all show that Europe The unique and multi-level endowments and abilities have already demonstrated Europe's gradual occupation of the global network governance system. The above examples of the "General Data Protection Regulation" (GDPR) cannot be ignored, the GDPR is still being constantly improved, which also illustrates the importance of open source data security and privacy.

1.2 Motivation

This project will use NetLogo (NetLogo will be introduced in detail in the second part) modeling methods to simulate the sharing and collaboration of open source data, and explore the impact of restrictions and protection measures on data sharing. The project is driven by the following considerations:

* The sharing of information and data is a major trend in today's society. The research on data sharing helps to understand the storage and dissemination of network information, and to understand how cross-country and cross-regional cooperation is achieved. This will also help me in the future. Learn how to better share and collaborate with others in your learning career.
* People’s reliance on the Internet is increasing, and more and more users tend to upload their personal data information (text messages, photos, videos, etc.) to the Internet for backup. Whether it is public or private, there are Risk of stealing and leakage. This project helps to understand the relevant standards and methods of data protection, and on this basis attempts to propose more secure methods.
* This project will also illustrate the impact of data protection on data dissemination, especially on speed and scope. Because we cannot blindly pursue stringent safety measures, we must also consider the practicality of these measures, and we must not lose sight of one another.
* There is not much work related to modeling research on pure data sharing and collaboration, especially when security is considered. Few studies have studied this issue by establishing simulation models. Therefore, the project analyzes and studies this problem from a modeling perspective, and encourages researchers to explore other ways to simulate the problem.

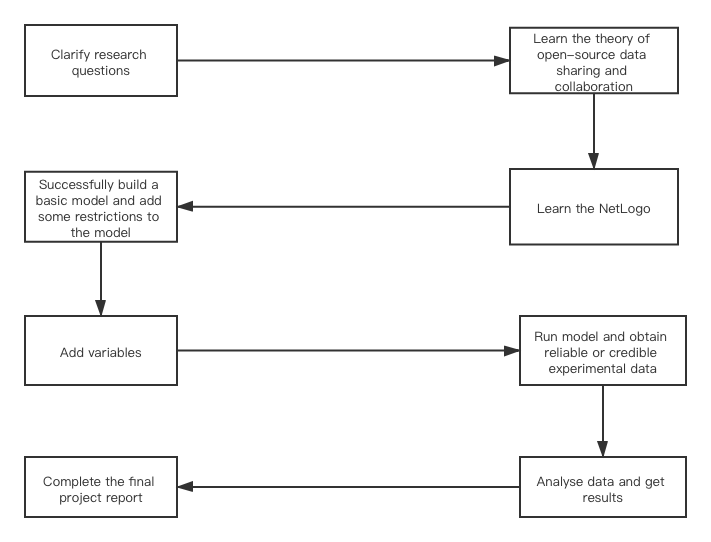
1.3 Research question

Modeling methods can help researchers better understand data sharing and collaboration. The focus of this project is to simulate this process by establishing a model, and to explore the impact of varying degrees of restrictions and protection measures on this process. Therefore, the following questions are also raised:

* How is open source data sharing and collaboration achieved?
* Is unlimited open source data safe?
* Is open source data with added protection measures absolutely safe?
* Will protection measures affect the speed and scope of data dissemination? If so, are the effects of different levels of protection measures the same?
* Can you get a more perfect model that performs in both aspects by continuously adjusting the model parameters?

1.4 Methodology

The realization of this project is divided into two parts. The first part is the research method to determine the problem. After browsing some relevant documents, we can learn that building models is a fast and accurate method for research on social issues. Therefore, this project will use modeling methods to simulate this problem and conduct research. The second part is the modeling process and realization. The modeling method of this project is as follows:



The method and tools of model realization will be explained in detail in the following chapters.