# 1. Introduction

Disclosing or sharing data to another party can have dual implications. On one hand, it facilitates collaboration between different organizations, drives innovation and research and improves decision-making. On the other hand, it also brings risks including privacy violations, unauthorized access, misuse of information, and theft of sensitive information. The consequences of such risks can range from minor financial losses to significant harm to individuals, organizations, and society. Thus, it is becoming more and more critical for organizations to weigh the possible risks before they disclose data. This article will use troubled family's data, which is highly sensitive, as an example through the Troubled Families Programme(TFP) which allows researchers access to the data, to analyze the risks involved and provide appropriate recommendations.

# 2. Background and Methodology

The Troubled Families Programme(TFP), which was launched in England in 2011 and ended in 2019, aimed to provide targeted support to families facing multiple and complex problems such as poverty, unemployment, poor educational achievement, and poor health(especially mental health). Many years of operation have given the TFP rich data, which was believed giving access to researchers who can gain greater insight and inform future policy and practice (e.g. Building Better Futures programme). However, getting more people involved in the project will inevitably lead to greater data risk.

Our risk analysis was based on filling in the Anonymisation Decision-Making Framework (ADF)(Elliot M, 2016), where Content 1-4 helped us to catch the changes in the data environment and the possible risks associated with this changing, which are summarised below as Scenario 1. In Content 6, we found that the TFP situation fell in the red zone, i.e. 'Essential ', indicating the importance of risk assessment and process controlling. Content 7 helps us with risk at the data level, which is summarised in Scenario 2. The specific ADF tables are attached in the appendix.

## 2.1 Assumptions

Before risk evaluation, we make the following assumptions.

- We assume that the current time point is 2022.
- We asscume the data we want to disclose to researchers is the data for each year of the WHOLE programme, i.e. 2011-2019, implying a DATA YEAR between 3 and 12 years.
- We assume that the intruder acts rationally and does not randomly engage in sabotage
- We assume that intruder's motivation is only two factors, financial and political.

# 3. Risk Analysis

Data risks primarily exist in two major dimensions: environmental, and data itself. In what follows, we present these two scenarios respectively.

## 3.1 Scenario One (Environment Risk)

Let's say there is a super-rich consortium who want to prove that the TFP is a wrong government

project to support his election. Hence, they hired hackers to access the TFP data directly and used it to generate statements in their favour. The hackers wanted to maximize their profits and get the most complete data with the least amount of effort, then they conducted an environmental analysis of the TFP data.

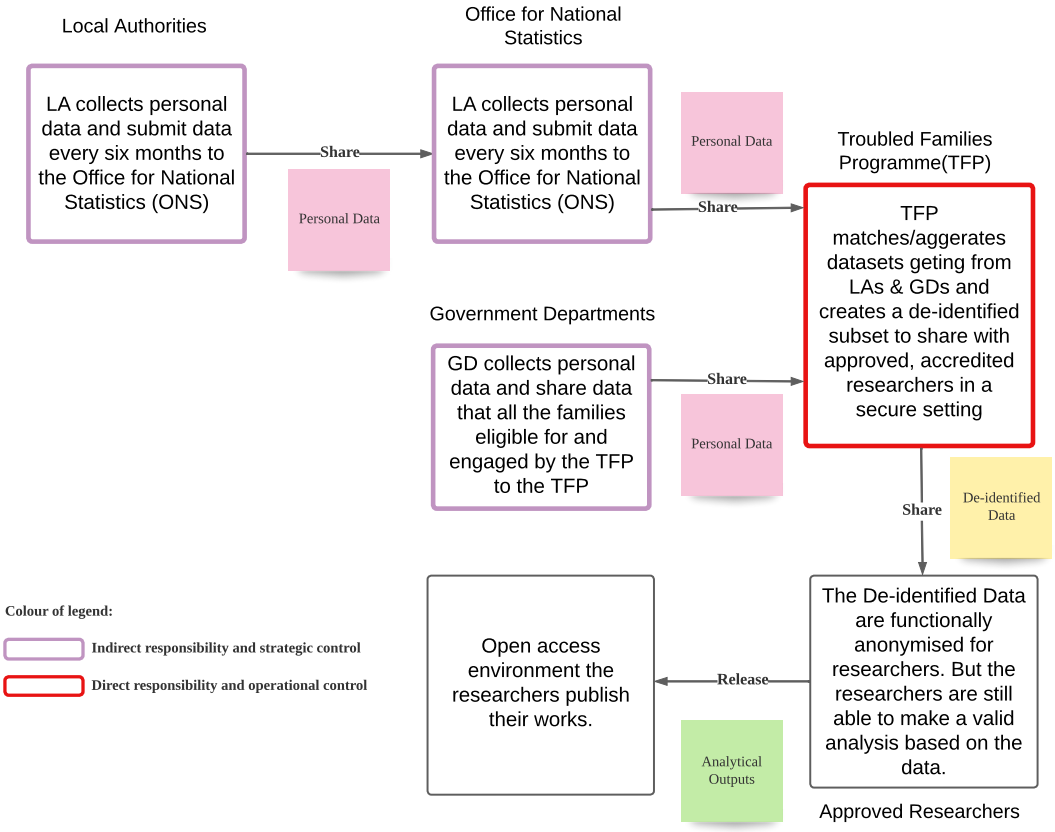


Figure 1. Data flow in different environments

TFP data is not fixed but flows through different environments. Government departments and local authorities[1][2][3][4]only provide a portion of the data, having indirect responsibility and strategic control. Access to the government departments and local authorities' datasets is strictly controlled, and only authorized individuals and organizations are permitted to use the data for approved purposes. TFP combine the information together, having direct responsibility and operational control on what data to choose. This means that before the TFP shares data with researchers, the data is closed and controlled.

As rational hackers, first they would prefer to access a closed database rather than an open data set. This is because closed databases typically contain more sensitive and personal information, such as names, addresses, and other identifying details, that can be used to reidentify individuals. In contrast, open data sets are often stripped of sensitive information or otherwise de-identified to minimize the

[1] Data Origin: combines administrative data from multiple government departments and local authorities. (On around 60,000 families data recorded)

[2] The Police National Computer (PNC) held by Ministry of Justice

[3] The National Pupil Database (NPD) held by Department for Education and the Work and Pensions Longitudinal Study (WPLS) held by Department for Work and Pensions

[4] Family Progress Data (FPD) by local authorities directly to DCLG

risk of reidentification. Thus, they will target the TFP environment and researchers as the first tier of attack.

As for the environments before TFP, although they were also closed data sets, they were dispersed and attacking each one individually would cost a lot. On the contrary, the data of TFP and researchers is already aggregated and specific to this programme. Furthermore, data owned by researchers has higher possibility to be attacked. TFP data was stored and managed by the Department for Communities and Local Government (DCLG), with more safety infrastructure to prevent intruder's attack. Once they share the data to researchers, these safety infrastructures cannot guarantee that the researchers' data will not be leaked. As more researchers involved in this programme, if one of their databases is vulnerable, hackers will be able to get the data they want. (e.g., Zimmer 2010)

## 3.2 Scenario Two (Data Risk)

In Scenario 1, we envisaged the possible risks posed by the environment, in this section, we analyse the risks of the data itself. We imagine that the intruder is a criminal gang that sells information. They sell illegally obtained personal data to insurance companies or fraud groups for profit. Their technical means are limited, but they have plenty of time and spend a lot of time analyzing publicly available data sets and reports. The main reason they chose to attack TFP is that this contains a lot of information about poor families, which can be used extensively for financial fraud, e.g., health insurance fraud. Health insurance fraudsters have been known to target poor families by submitting false claims for medical treatments or procedures that were never performed.

### 3.2.1 High risk direct identifier

Through the report and metadata for potential TFP dataset released by TFP, we have marked the following sensitive attributes. Among them, Name is a direct identifier, although there are people with the same name, but with other information, we can easily identify a person.

Table 1. Risk attributes

| Variable Type | Variable Meaning |
|---|---|
| Direct | Local authority name |
| Indirect | The gender of the child |
| Indirect | Child's year of birth |
| Possibly | The number of people in that child's family at the time of their birth |
| Possibly | The child's age when their family member had their first mental health and/or drug or alcohol episode |
| Possibly | If lone parent families |
| Possibly | If family with an individual with any mental health issue |
| Possibly | If Family with an individual dependent on drugs or alcohol |
| Possibly | If families who have been involved in a domestic abuse incident |

### 3.2.2 Longitudinal data

In Table 1, the green attributes indicate potential identifiers. Some belongs to the longitudinal attribute, which changing over time. For example, by concatenating the data with the previous year, intruders can locate information on the new birth family and, combined with information on births in the local ward, may be able to obtain information on a local trouble family's information.

### 3.2.3 Hierachical data

Another class belongs to hierarchical attribute. This type of data may appear in multiple data sets, and if an intruder merges two tables with similar variables, they can find the overlapping portion. For example, if a troubled family has a higher risk of mental illness and drug abusing, then linking the TFP to the local hospital discharge record will allow for reidentification. (Ohm, 2010)

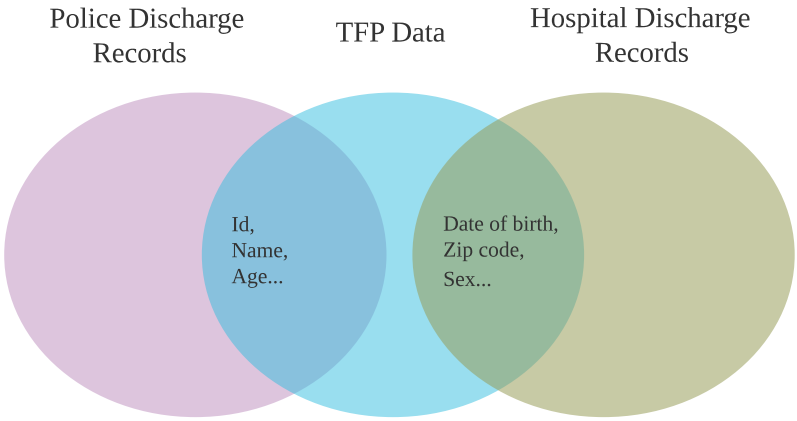


Figure 2. Hypothetical example of reidentification attack of the Troubled Families Programme

## 3.3 Risk Mitigation

From the discussion of the two scenarios, we conclude that in the case of TFP situation, the main threats come from the researchers, and data that is chosen to be disclosed. To mitigate risks, some steps can be taken include:

- Cryptography and environmental security: Hire network security experts, upgrade database security systems; Secondary passwords permissions for researchers. They are required passwords to access the database, and another certificate to open the data. (Reiter and Kinney 2011)
- Modification to Data: De-identify the data when we share it with the researchers, removing the direct identifier, differencing the data or applying k- anonymity (Sweeney 2002b), and turning a specific attribute, such as age, into an interval.
- Publicity: keep a low profile when bringing in researchers. Low-profile research is less likely to be noticed by these opponents and may be a less attractive target.
- Ethical Appeal: Each participant and researcher should read and sign the statement defining how he or she would use the data responsibly.
- Regular security assessments: Conducting regular security assessments and penetration testing can help identify and address vulnerabilities in the database.

# 4. Conclusion

This case study analysed the possible risks among disclosing the TFP data to researchers in

environmental and data itself aspects. When deciding share data to another party, we should take the risks it may occur. There are many approaches in different areas can be used to mitigate the risks. We hope this case study showing the possible that safe and responsible data access will be ensured through collaboration from different sectors to improve the technical, legal, ethical, and social frameworks.

## Reference

Mackey E, Elliot M, O' Hara K. The anonymisation decision-making framework[M]. UKAN publications, 2016.

Ohm, Paul.2010. "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization." UCLA Law Review 57(6):1701.

Kinney, Satkartar K., Jerome P. Reiter, Arnold P. Reznek, Javier Miranda, Ron S. Jarmin, and John M. Abowd. 2011. "Towards Unrestricted Public Use Business Microdata: The Synthetic Longitudinal Business Database." International Statistical Review 79(3):362–84.

Sweeney, Latanya. 2002b. "k-Anonymity: A Model for Protecting Privacy." International Journal of Uncertainty, Fuzziness and Knowledge-based Systems 10(05):557–70.

Zimmer, Michael. 2010. " 'But the Data Is Already Public' : On the Ethics of Research in Facebook." Ethics and Information Technology 12(4):313–25.