The establishment of a Central Credit Register at the Bank of Israel and its Statistical Disclosure Control processes

Ariel Mantzura
Bank of Israel

August 31, 2018

Credit Data Law - Objectives

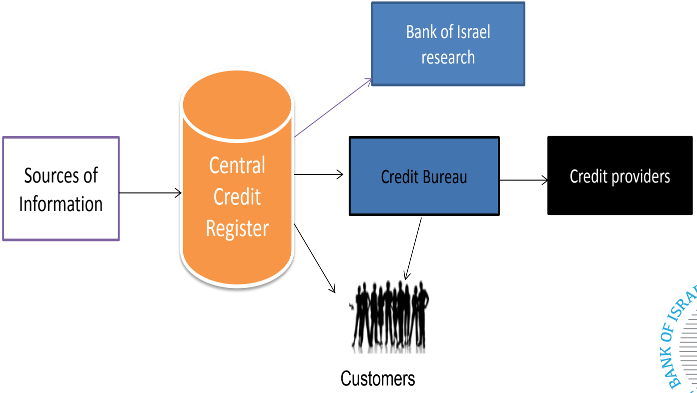
The objective of this law is to establish an overall arrangement for sharing credit data...for the following purposes:

- Enhancing competition in the retail credit market.
- Expanding access to credit.
- Reducing of discrimination in the granting of credit and of economic gaps.
- Creating an anonymous database for use by the Bank of Israel in carrying out its functions.

Backround - Protection of Privacy Law

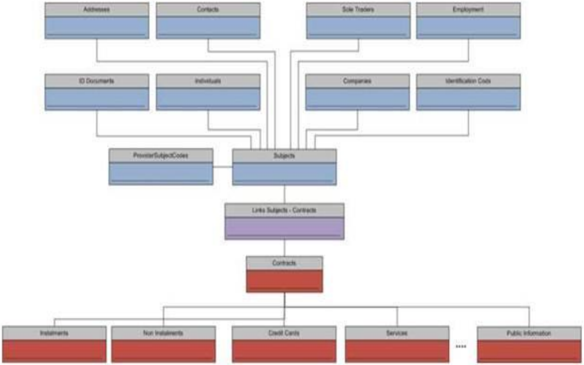
- In order to allow access to such information within the organization or outside it, the Protection of Privacy Law requires that the confidentiality of the information be maintained, as the information relates to individual persons.
- In addition, the law requires that the commercial confidentiality of business entities be maintained, a complex task, particularly when dealing with financial information that is sometimes characterized by high concentration.

Backround - Central Credit Register



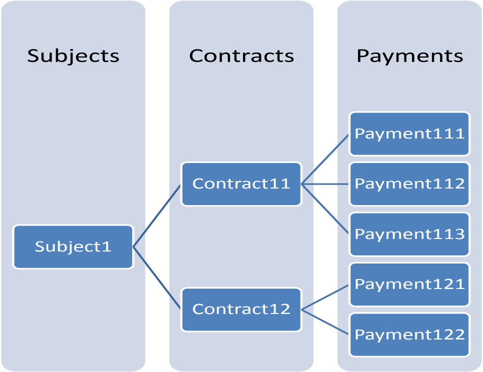
Backround - Credit register structure

Linked tables structure

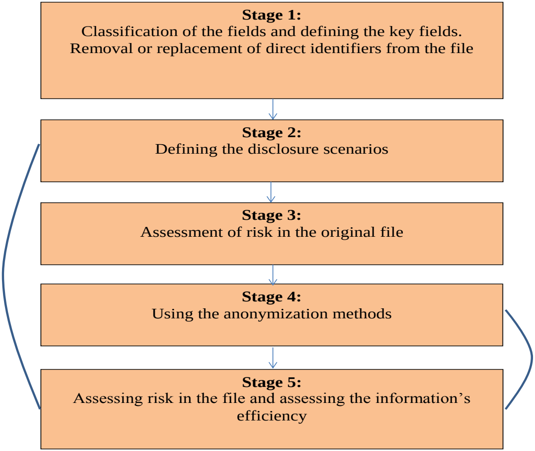


Backround - Credit register structure

Flat structure



Flow chart of anonymization process



Stage 1:
Classification of the fields and defining the key fields.
Removal or replacement of direct identifiers from the file

Stage 2:
Defining the disclosure scenarios

Stage 3:
Assessment of risk in the original file

Stage 4:
Using the anonymization methods

Stage 5:
Assessing risk in the file and assessing the information’s efficiency

Defining disclosure scenerios

- Disclosure scenarios are a group of assumptions that describe how a user, or another person exposed to the file, can expose information on individuals from within the file.
- For instance: A user can cross-reference the information from the file with other information he has through a number of common characteristics.
- The disclosure scenario can for the most part be summed up by determining groups of key fields through which information in the file can be cross-referenced with other external information.

Defining disclosure scenerios

- Setting disclosure scenarios is necessary to the anonymization process, since we are trying to protect the information from them.
- The assessment of the level of risk of information disclosure is also dependent on setting these scenarios.
- The disclosure scenarios are determined with the help of experts in the relevant content worlds.

Defining disclosure scenerios

Database that can be cross-referenced	**Type of data in credit register that can be cross referenced**	**Who has access to both databases?**
Personal information	• Direct identifiers • Exact numeric values	Credit register users
Mortgage file data	•	•
Employees file	•	•
Real-estate transactions file	•	•

Defining disclosure scenerios

- The disclosure scenarios can be less or more severe than the objective information disclosure possibilities.
- The disclosure policy depends on how the data are used, the purpose of the use, the identity of the users, the severity of the damage inherent in disclosure, and so forth.

Defining disclosure scenerios

- It is common to distinguish between **scientific use files** and **public use files**.
- **scientific use files (SUF)** are used by researchers under contract, subject to permissions and restrictions such as working within a physical research room or a virtual research room through remote access.
- **public use files (PUF)** have no restriction or control. The policy regarding the information files issued to the public is generally very strict.

Assessing the risk of disclosure in the file

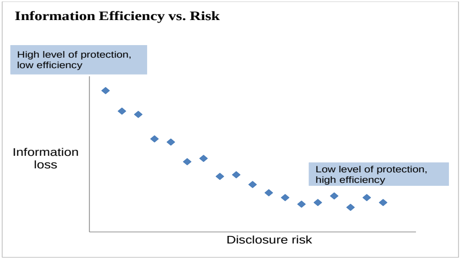
- There are two common requirements:
- **K-anonymity** - a requirement that in each combination of categorical key fields in groups that are defined in the disclosure scenario, there shall be at least k records with the same combination.
- **l - diversity** - The l-diversity requirement is that in all possible combinations there should be at least l different values.

Information efficiency vs minimizing risk

- The objective of the anonymization process is to make a protected file of data accessible so that it embodies a low risk of identification of the individuals.
- At the same time, subject to that limitation, it maintains maximum information in the file.
- There is a tradeoff between the level of information protection and its usability.
- The higher the level of protection, the greater the information loss.

Information efficiency vs minimizing risk

Figure: Risk Utility Map



Thank you!