2(B) RISK-VALUE ANALYSIS OF DATA PROCESSING[1]

The complexity of balancing utility and privacy in open data means that there is no "correct"decision for any dataset: processing or publicating data carries benefits for the public as well as potential risks to individual privacy.

Therefore, it is recommended to continually evaluate the risks and benefits of the processing of certain data sets. This risk-value analysis is one systematic way of doing so. The tool can be used to assess data processing, for example before the conception of a digital project, or for the evaluation of databases that are intended to be opened up in the context of open data policies. Keep in mind that this analysis is aimed at guiding a decision but its result should not be taken as an absolute.

Phase 1: Evaluate the value of the data processing for citizens

Impact \ Probability	Low	Moderate	High
Low	L	L	M
Moderate	L	M	H
High	M	H	H

IMPACT = The potential positive benefit for citizens if the data processing takes place.

PROBABILITY = The likelihood that citizens will actually gain the positive benefits of the data processing.

Phase 2: Evaluate the risk in four steps

Table 2: Which expectations of privacy do citizens have?

Low	i.e. data is about public matters; the data is disclosed for the identical reason to the one it was collected for; notice of disclosure was given at the time of collection.
Moderate	i.e. data could be considered public or private; the data is disclosed for a similar or tangential context to which it was collected; no notice was given at the time of collection.
High	i.e. the data refers to the private life of a person; the data is disclosed for a context unrelated to the one for which it was collected; Notice was given not to disclose the information at the time of collection.

Table 3: What could be the consequences in case an individual is re-identified in the data?

Of no concern	Individual - emotional damage or physical harm
Minor	Individual - reputation damage Individual - financial damage Organization - lack of compliance Organization - reputational damage Organization - financial damage
Moderate	Society - exclusion, marginalization or discrimination of groups
Major	Society - "targeting" of groups

Table 4 Evaluate the potential risk (negative impact) based on the results of table 2 and table 3

Expectations of citizens \ Consequences	Of no concern	Minor	Moderate	Major
Low	L	L	M	H
Moderate	L	M	H	H
High	M	H	A	A

Table 5 How likely is it that the data will be misused?

Impact \ Probability	Low	Moderate	High
Low	L	L	M
Moderate	L	M	H
High	M	H	H

PROBABILITY = Likelihood that the risk will occur

IMPACT = Result of table 4

Phase 3: Risk-Value Ratio

Table 6: What is the risk-value ratio?

Risk \ Value	Low	Moderate	High
Low	L	L	M
Moderate	L	M	H
High	M	H	H

Given the result of the analisis, should you collect / process / open the data?

- ☐ YES
- ☐ NO

If the response is "no" don't proceed with the following question

VALUE = Result of table 1
RISK = Result of table 5

Phase 4: Risk Mitigation

If the response is "yes", what could be potential mitigation mechanisms to decrease the risk (multiple options possible):

- ☐ Safe storage
- ☐ Minimize the collection of personal data
- ☐ Access controls (Who can access what data)
- ☐ Integrate audit trails
- ☐ (Document what happens when in the system)
- ☐ Delete fields
- ☐ Delete records
- ☐ Add noise to the data
- ☐ Generalize data
- ☐ Anonymize data
- ☐ Backup files
- ☐ Others: _____

Table 7 After applying these mitigation measures how would you reevaluate the risk-value ratio?

Risk \ Value	Low	Moderate	High
Low	L	L	M
Moderate	L	M	H
High	M	H	H

Phase 5: Final Decision

Write down the final decision about collecting/processing/opening up the data.

_____

_____

_____

_____

_____

This tool is based on the document "San Francisco: „Open Data Release Toolkit".