# A framework for quantifying the Adaptive Cycle of Resilience

**Author**
Timon Brouwer
TimonBrouwer19@gmail.com
University of Amsterdam

**Internal supervision**
Toon Abcouwer
A.W.Abcouwer@uva.nl
Otte-Pieter Banga
ottepieter.banga@gmail.com

**External supervision**
David Macro
David@dataim.nl
DataIM

## 1. Introduction

This document describes the context and the planning of the data collection performed for a Master's thesis done at DataIM. DataIM is a survey company that facilitates the conducting of surveys and provides tooling and services to analyze these surveys.

In this research, an attempt will be made to quantify resilience and show its importance to organizations and their stakeholders. This will be done using the Adaptive Cycle of Resilience (ACoR) model as a framework for resilience and a survey design with the intent to measure concepts from the ACoR model. In order to capture and rank the importance of these concepts, the survey will make use of a combination of likert-scale questions and open-ended questions. Often the answers to these open-ended questions are ignored (Schonlau et al. 2021), so to properly analyze this textual data, an NLP technique called Latent Dirichlet Allocation (LDA) will be used to extract topics and their importance from the responses. Ultimately, the research delivers a dashboard where the quantified resilience can be presented and compared with other measurements along with proper validation.

**Research question**

"How can the key concepts of resilience, based on the ACoR model, be quantified using surveys and Structural Topic Modelling in small and medium-sized organizations?"

**Sub questions**

- What are the key concepts of resilience that can be measured in a survey?
- What are key concepts of resilience measurable with surveys?
- How can LDA be used in open-ended answers of varying lengths and how can they deepen understanding of the key concepts?
- How can organizations be supported in making their initial steps from an Equilibrium state to a new combinations state?

## 2. Survey structure

The structure of the survey is based on the work of Stephenson et al. (2010) and Goncalves et al. (2019), who have performed similar research. In the research of Stephenson et al. (2010), a survey was conducted based on four main concepts of resilience and a multitude of sub-concepts (appendix A). The survey consisted of 92 likert-scale statements from 1 to 4 and a "I don't know" option. These statements where distributed over the four main concepts of resilience they identified. Goncalves et al. (2019) iterated further on this idea, changing the amount of questions to 13, keeping the same structure (appendix B).

The survey of this thesis will use a slightly adjusted structure of these researches. Instead of a four-point likert scale, a six-point likert scale will be used to open the possibility for more kinds of validation tests. In addition, the "I don't know" option will not be included so responses will not have to be deleted for not containing enough usable information. Lastly, an open-ended question will be added for each key concept of resilience as identified by Stephenson et al. (2010). In this open question, the respondent described their perspective on the key concept. At the end of the survey, the respondent will be asked if they can be contacted for an interview to validate the results of the survey. This field will be optional. All this leads to the following structure where all but the last point are mandatory:

1. **Front page** - description of the survey
2. **Situation Awareness** – 1 open question, 3 likert-scale questions
3. **Management of keystone vulnerabilities** - 1 open question, 3 likert-scale questions
4. **Adaptive capacity** - 1 open question, 3 likert-scale questions
5. **Resilience ethos** - 1 open question, 3 likert-scale questions
6. **Contact** – optional information field for contact for a validation interview

## 3. Data collection

**Sampling methods**

To find respondents for the survey, one or a combination of sample methods will be used. Both sample methods will have different quota and will lead to a slightly different analysis of the survey. The choice of method will be based on whether the first proposed method is feasible and the costs of the second proposed method.

First, a multistage sampling method will be explored. In this method a combination of convenience sampling and stratified sampling will be used. Convenience sampling will be used to find organizations that are willing to work together on a pilot on measuring resilience. It is preferred to conduct surveys with organizational entities to be able to draw organization-wide conclusions based on the liker-scale questions and the open-questions in the survey. An LDA analysis for example might lead to a specific challenge or disruption the organization is facing or about to face. After convenience sampling, stratified sampling will be used among the staff of the organization. The main groups that people will be divided on are management and workfloor staff. Dividing the staff into these groups might lead to the exposition of differences in resilience perspective, which can be further analyzed. The quota of this method is to conduct surveys together with at least 2 companies in a similar sector, of at least 50 staff members. The reason for this is to be able to compare the measurements with each other for further analysis.

Second, a simple random sampling method will be explored. In this method, a third party service (PanelClix) will be used to acquire responses from several organizations in the Netherlands. With this method, a quota of 500 responses will be met. The benefit of this method is that the population characteristics can be controlled (e.g. an equal amount of men and women) and that a higher number of respondents can be obtained. There are however multiple disadvantages attached to this method as well. First the analysis of the survey won't be able to draw conclusions for specific organizations, but for the population as a whole. Second, challenges and disruptions particular to an individual organization will not be detected, since the respondents are not clustered. Third, a cost is attached per respondent and on specifying population characteristics like location and age. Therefor the first method will take preference.

**Limiting non-response**

Limiting non-response will only be applicable for the first method introduced in the previous section. For the simple random sampling method, PanelClix will be responsible for collecting enough respondents based on the agreement made with them. In case the first method is used, clear agreements will be made with the partnering organizations on the response rate of the survey where at least 50% of the staff will fill in the survey. This number is reached based on previous projects in the organisation.

Apart from the initial response rate, there is also a group of respondents that will abort during the survey. The size of this group is longer surveys, harder questions and technical issues during the survey. For the first factor, the survey has been limited to 20 questions, as opposed to similar research that is performed with 90 questions (Stephenson et al., 2010). The second factor will be combatted by only providing likert-scale questions and open-questions and not choosing for unordinary question types. As for the third factor, no routing in the survey will be used as faults in the routing are one of the main reasons for surveys not properly functioning. DataIM has facilitated hundreds of surveys, so other causes of technical failure are not to be expected.

**Data storage**

The responses of the survey will be stored on a company-owned SQL server. In this server, the metadata from the survey will be separated from the main answers to the questions. The responses of the survey will be anonymized where needed. Direct queries can be made from Python, or an export to Excel or SPSS can be made to start analyzing the data.

## 4. EDA

When the responses are finalized, the data will have to be cleaned and an EDA will be performed. In this section, this process will be described.

**Data format**

The data of the survey will have a format that is widely used in DataIM. In the survey structure there are likert-scale questions and open-questions. Each open-ended question will result in a single column. The likert-scale questions will be inserted as a grid, which leads to a column for each answer of the question (figure 1).

| Q4A1 | Q4A2 | Q4A3 | Q4A4 | Q4A5 | Q4A6 | Q4A6_open |
|---|---|---|---|---|---|---|
| False | False | False | False | True | False | |
| False | False | False | False | False | True | Via verzekeraar van buurvrouw |
| False | True | False | False | False | False | |
| False | False | False | False | True | False | |
| False | False | False | True | False | False | |

*Figure 1. Data structure example*

Furthermore, the data format of the survey consists of initial values and labels for these values in the form of a mask. The labels can be used to replace the question codes with the actual question text. In the case a different question type is chosen after all, the labels for the values of the data might have to be used to be able to analyze the survey.

Lastly, in addition to the answers for each question, metadata of the survey is stored (figure 2). This data is import to recognize any aborts during the filling of the survey. Aborts will generally have to be removed, unless it is determined that the data is still valuable. The metadata is also used to see when people close the survey and continue at another point in time. When that is the case, the same 'CODE' will be used, but the response will be under different ID's. These rows will have to be merged.

| | ID | CODE | STARTDATE | ENDDATE | STATE |
|---|---|---|---|---|---|
| 0 | 1.615964e+09 | TobefaijQYQEb | 2021-03-17 10:12:52.530000640 | 2021-03-17 10:15:14.170000128 | Complete |
| 1 | 1.615967e+09 | TIcvgIIzMKGuc | 2021-03-17 09:30:52.139999488 | 2021-03-17 09:36:23.746999808 | Complete |
| 2 | 1.615968e+09 | TUfQvBIIKhy | 2021-03-17 09:37:39.382999296 | 2021-03-17 09:46:38.113000960 | Complete |

*Figure 2. Metadata of the survey.*

**Cleaning**

The data will be cleaned and analyzed in Python with the Pandas library. Missing values shouldn't exist, since all questions are mandatory. After loading in the data, the following steps will be taken

1. Replace values with their labels
2. Remove aborts and merge different sessions of the same user
3. Transform the likert-scale questions so they give a score per question instead of True / False per answer
4. Anonymize open questions (in case personal details have been filled in like names of persons and organizations)
5. Preprocess the open-questions to perform an LDA

While the likert-scale questions can be used after taking the steps above, open-questions have to undergo more steps to be cleaned so they can be analyzed with an LDA:

1. Clean unwanted characters from the text and empty responses
2. Anonymize the text (no references to persons or organizations)
3. Tokenize
4. Remove stopwords
5. Stemming / lemmatization

**Exploration**

For the exploration of the data, first the population will be checked. This means checking how many respondents there are per company and the distribution of managers and workfloor staff. Then correlations between the likert-scale questions will be checked.

For the open-questions, word clouds will be made to see what topics people have been writing about.

## Github

https://github.com/IIVolumeII/Quantifying-resilience

## References


[1] Gonçalves, L., Navarro, J. B., & Sala, R. (2019). Spanish validation of the Benchmark Resilience Tool (short-form version) to evaluate organisational resilience. Safety science, 111, 94-101.

[2] Schonlau, M., Gweon, H., & Wenemark, M. (2021). Automatic classification of open-ended questions: check-all-that-apply questions. Social Science Computer Review, 39(4), 562-572.

[3] Stephenson, A., Vargo, J., & Seville, E. (2010). Measuring and comparing organisational resilience in Auckland. *Australian Journal of Emergency Management*, *25*(2), 27-32.

# Appendix

Appendix A

| **TABLE 1:** Updated Dimensions and Indicators of Organisational Resilience (Adapted from McManus, et al., 2007, p. 20) | | | | | |
|---|---|---|---|---|---|
| **Resilience Ethos** | | | | | |
| $RE_1$ | Commitment to Resilience | | | | |
| $RE_2$ | Network Perspective | | | | |
| **Organisational Resilience Factors** | | | | | |
| **Situation Awareness** | | **Management of Keystone Vulnerabilities** | | **Adaptive Capacity** | |
| $SA_1$ | Roles & Responsibilities | $KV_1$ | Planning Strategies | $AC_1$ | Silo Mentality |
| $SA_2$ | Understanding & Analysis of Hazards & Consequences | $KV_2$ | Participation in Exercises | $AC_2$ | Communications & Relationships |
| $SA_3$ | Connectivity Awareness | $KV_3$ | Capability & Capacity of Internal Resources | $AC_3$ | Strategic Vision & Outcome Expectancy |
| $SA_4$ | Insurance Awareness | $KV_4$ | Capability & Capacity of External Resources | $AC_4$ | Information & Knowledge |
| $SA_5$ | Recovery Priorities | $KV_5$ | Organisational Connectivity | $AC_5$ | Leadership, Management & Governance Structures |
| $SA_6$ | Internal & External Situation Monitoring & Reporting | $KV_6$ | Robust Processes for Identifying & Analysing Vulnerabilities | $AC_6$ | Innovation & Creativity |
| $SA_7$ | Informed Decision Making | $KV_7$ | Staff Engagement & Involvement | $AC_7$ | Devolved & Responsive Decision Making |

Appendix B

| Factor | Indicator | Item |
|---|---|---|
| Planning | P1 | We are mindful of how a crisis could affect us |
| | P2 | We believe emergency plans must be practised and tested to be effective |
| | P3 | We are able to shift rapidly from business-as-usual to respond to crises |
| | P4 | We build relationships with organisations we might have to work with in a crisis |
| | P5 | Our priorities for recovery would provide direction for staff in a crisis |
| Adaptive Capacity | AC1 | There is a sense of teamwork and camaraderie in our organisation |
| | AC2 | Our organisation maintains sufficient resources to absorb some unexpected change |
| | AC3 | People in our organisation "own" a problem until it is resolved |
| | AC4 | Staff have the information and knowledge they need to respond to unexpected problems |
| | AC5 | Managers in our organisation lead by example |
| | AC6 | Staff are rewarded for "thinking outside the box" |
| | AC7 | Our organisation can make tough decisions quickly |
| | AC8 | Managers actively listen for problems |