

Contents

Acknowledgments – Agraïments – Agradecimientos – Remerciements	iv
Abstract.....	vi
Resum.....	viii
Resumen	x
Contents.....	xiii
Chapter 1. Introduction	1
Chapter 2. Literature review.....	6
Chapter 3. Conceptual framework and theoretical strategy	25
3.1. Definition of <i>data</i> , <i>primary data</i> , <i>secondary data</i> and <i>reuse of data</i>	25
3.2. Rational choice theory, bounded rationality and procedural rationality	31
3.3. A model of the scientific actor’s behavior and decision-making: the <i>bounded individual horizon</i> (BIH) model	38
Chapter 4. Methodology and methods	42
4.1. Research questions and an approach to answer them.....	44
4.2. Researchers’ decisions based on the BIH model when reusing data: the data-reuse mechanism.....	46
The researcher’s structure and causal powers and liabilities.....	51
Condition C1 – The researcher knows that secondary data exist	52
Condition C2 – Secondary data are obtained	53
Condition C3 – Particular secondary data are an initial satisficing option	53
Condition C4 – The idea of collecting particular primary data is not an initial satisficing option.....	54

Condition C5 – An expected scientific contribution exists and the researcher finds its potential rewards initially satisficing	54
Potential events of the data-reuse mechanism	55
4.3. Justification of a multi-case study approach	59
4.4. The search process of case studies	65
4.5. Description of data collecting methods and instruments.....	67
The interview	68
Authors’ publications and their complementary role	69
Visual representation of each participant’s data reuse process	70
Follow-up messages	71
4.6. Data analysis methods.....	71
Within-case analysis.....	73
Cross-case analysis.....	74
4.7. A diachronic process of data collection and data analysis	74
4.8. Ethics protocols and data sharing agreements.....	79
4.9. A small exercise on reflexivity.....	79
Chapter 5. General overview of cases and data sources collected for each case	82
5.1. Case studies reusing <i>released data</i>	85
5.1.1. Collected empirical data and collection dates in case study #1 (GTEX data repository)	88
5.1.2. Collected empirical data and collection dates in case study #2 (GEO Profiles repository)	89
5.1.3. Collected empirical data and collection dates in case study #3 (TCGA data repository)	90
5.1.4. Collected empirical data and collection dates in case study #4 (GEO Profiles and TCGA repositories).....	91
5.2. Case studies reusing <i>stewarded data</i>	92
5.2.1. Collected empirical data and collection dates in case study #5 (BORN Ontario data & ICES data)	95
5.2.2. Collected empirical data and collection dates in case study #6 (BORN Ontario data)	96
5.2.3. Collected empirical data and collection dates in case study #7 (BORN Ontario data)	97
5.3. Case studies reusing <i>proprietary data</i>	97
5.3.1. Collected empirical data and collection dates in case study #8 (IPD MA)	100
5.3.2. Collected empirical data and collection dates in case study #9 (IPD NMA)	101
5.3.3. Collected empirical data and collection dates in case study #10 (IPD NMA)	102

Appendix.....	216
Annex 1. Interview script vs3 IAB Nov 2016.....	217
Annex 2. Data collection instruments and dates. Case study #1	222
Annex 3. Data collection instruments and dates. Case study #2	223
Annex 4. Data collection instruments and dates. Case study #3	224
Annex 5. Data collection instruments and dates. Case study #4.....	225
Annex 6. Data collection instruments and dates. Case study #5	226
Annex 7. Data collection instruments and dates. Case study #6.....	227
Annex 8. Data collection instruments and dates. Case study #7	228
Annex 9. Data collection instruments and dates. Case study #8.....	229
Annex 10. Data collection instruments and dates. Case study #9	230
Annex 11. Data collection instruments and dates. Case study #10	231
Annex 12. Workflow diagram of data reuse process. Case study #1	232
Annex 13. Workflow diagram of data reuse process. Case study #2	233
Annex 14. Workflow diagram of data reuse process. Case study #4	234
Annex 15. Situating the reuse of data within a larger research enquiry. Case study #4	235
Annex 16. Workflow diagram of data reuse process. Case study #5	236
Annex 17. Workflow diagram of data reuse process. Case study #6	237
Annex 18. Situating the reuse of data within a larger research project. Case study #6.....	238
Annex 19. Workflow diagram of data reuse process. Case study #7	239
Annex 20. Workflow diagram of data reuse process. Case study #8	240
Annex 21. Workflow diagram of data reuse process. Case study #9	241
Annex 22. Workflow diagram of data reuse process. Case study #10	242
Annex 23. Structure of the data-reuse mechanism.....	243
Annex 24. Literature about factors affecting the reuse of data by IS scholars.....	244
Annex 25. Process-tracing of the data-reuse mechanism when data are reused as the only evidence of scientific claims	250
Annex 26. Process-tracing of the data-reuse mechanism when secondary data are used to support scientific claims done with primary data.....	251
Annex 27. Process-tracing of the data-reuse mechanism when condition C5 is not met in time 2.....	252

List of Tables

Table 1 - Two possible initial combinations A and B of conditions of the data-reuse mechanism, and their respective hypothesized outcomes.....	58
Table 2 - Key of the images and symbols used in the visual representation of data collection instruments and dates.....	84
Table 3 - Variability of case studies reusing "released data"	87
Table 4 - Variability of case studies reusing "stewarded data"	94
Table 5 - Variability of case studies reusing "proprietary data"	99
Table 6 - Summary of outcome and conditions of findings in case study #6	176
Table 7 - Summary of outcomes and conditions of case studies #4 and #5.....	176
Table 8 - Summary of outcome and conditions of case studies #8, #9 and #10	177
Table 9 - Summary of outcome and conditions of case study #7	178
Table 10 - Summary of outcome and conditions of case study #1	179
Table 11 - Summary of outcome and conditions of case studies #2 and #3	180
Table 12 - Summary of analysis of the ten case studies	181
Table 13 - Several combination of conditions C3, C4 and C5 that do not lead to the reuse of data.	189

List of Figures

Figure 1 - Definition of data in this dissertation	29
Figure 2 – Visual representation of the definition of reuse of data or use of secondary data	30
Figure 3 - Source: Figure 7 – The structures of causal explanation (Sayer, 2010, p. 74)	47
Figure 4 - The data-reuse mechanism and its structure and potential events.....	57
Figure 5 - Four case studies of reuse of “released data”. Same researcher in case studies #1, #2, #3 reuses data from three different repositories (A, B, C). Researcher in case study #4 reuses data from repositories B and C.....	86
Figure 6 - Data collection instruments and dates. Case study #1	88
Figure 7 - Data collection instruments and dates. Case study #2.....	89
Figure 8 - Data collection instruments and dates. Case study #3.....	90
Figure 9 - Data collection instruments and dates. Case study #4.....	91
Figure 10 - Three case studies of reuse of "stewarded data". Three different researchers (case studies #5, #6, #7) reuse data from the same repository (BORN Ontario). One researcher (case study #5) reuses also data from ICES repository	93
Figure 11 - Data collection instruments and dates. Case study #5.....	95
Figure 12 - Data collection instruments and dates. Case study #6.....	96
Figure 13 - Data collection instruments and dates. Case study #7.....	97
Figure 14 – Three cases studies of reuse of “proprietary data”. Three different researchers, rather a research team, (case studies #8, #9, #10) reuse individual participant data (IPD) from different data sets in three different health problems	98
Figure 15 - Data collection instruments and dates. Case study #8.....	100
Figure 16 - Data collection instruments and dates. Case study #9.....	101
Figure 17 - Data collection instruments and dates. Case study #10.....	102
Figure 18 - Workflow diagram of the data reuse process of case study #1	107
Figure 19 - Workflow diagram of the data reuse process of case study #3	115

Figure 20 - Workflow diagram of the data reuse process of case study #4 123

Figure 21 - Situating the reuse of data within a larger research inquiry. Case study #4 125

Figure 22 - Dimensions, elements, and sub-elements of BORN’s Data Quality Framework. Source: BORN’s DQF 130

Figure 23 - Workflow diagram of the data reuse process of case study #5 134

Figure 24 - Situating the reuse of data within a larger research project. Case study #6 142

Figure 25 - Workflow diagram of the data reuse process of case study #6 143

Figure 26 - Workflow diagram of the data reuse process of case study #7 150

Figure 27 - Workflow diagram of the data reuse process of case study #8 157

Figure 28 - Workflow diagram of the data reuse process of case study #9 164

Figure 29 - Number of eligible studies and participants. Case study #9. Source: (Welch et al., 2019, p. 12)..... 166

Figure 30 - Workflow diagram of the data reuse process of case study #10 171

Figure 31 – Process-tracing of the data-reuse mechanism when data are reused as the only evidence of scientific claims 183

Figure 32 - Process-tracing of the data-reuse mechanism when secondary data are used to support scientific claims done with primary data. 183

Figure 33 - Process-tracing of the data-reuse mechanism when condition C5 is not met in time 2 .186

Figure 34 - Association between high effort and the goal of making a scientific contribution in the ten case studies..... 192