

(DRAFT)

**Open Innovation Learning:
Theory building on
open sourcing while private sourcing**

David Ing

April 28, 2017

Abstract

Organizations embark on *open innovation* initiatives to sweep in external knowledge, practices and resources in cooperation with partners. This contrasts to the mainstream *private innovation* approach of in-house research and development sponsored solely by an incorporated funder, with intellectual property protected by copyright. Few organizations simultaneously engage in both approaches, within and across the levels of programs, projects and individuals. How does *learning* occur in such an organization – and the communities of members within the organization – in both cumulative and distributed ways?

The *open innovation learning* exhibited by IBM in the decade of 2001-2011 provides a foundation for building both descriptive theories and normative theories. Legal protocols for *open source* licensing began in 1998, and "open innovation" became popular in the business press from 2003. At the beginning of the 2001-2011 period, a behaviour of *open sourcing* by commercial enterprises departed from traditional of *private sourcing* that presumes trade secrets for competitive advantage that maintaining economic viability. After a decade, *Open Sourcing while Private Sourcing (OSwPS)* had been demonstrated as a successful of doing business at IBM, and had also become adopted by other companies and institutions.

The primary method employed to appreciate the phenomenon of OSwPS is multiparadigm inquiry. Theories are developed inductively from seven case studies in a process orientation observing events, activities and choices ordered over time. Three descriptive theories have been built in parallel perspectives based on contrasting philosophies. Pursuit of a normative theory subsequently led to the proposal of additional hypotheses.

Emerging theories of open innovation learning challenge a presumption that commercial and non-commercial interests are incompatible. Open sourcing *while* private sourcing is a demonstrable way of conducting a viable business.

Preface

Preface to be written

Toronto, Canada, April 28, 2017,

David Ing

Contents

- Preface** **iii**

- List of Figures** **v**

- List of Tables** **vi**

- 1. Research approach: inductive from case studies** **1**
 - 1.1 Data: The history of open sourcing while private sourcing is observed as events, activities and choices ordered over time 2
 - 1.1.1 Process data: Over a decade, ways that open sourcing does and doesn't work with private sourcing were discovered 2
 - 1.1.2 Multilevel data: open sourcing while private sourcing coevolved for individuals, teams, corporations and non-profits 3
 - 1.2 Analysis: In hindsight, processual abstractions of evolutionary stages of open sourcing while private sourcing can be constructed 4

List of Figures

1.1	Considerations in research approach by chapter	1
-----	--	---

List of Tables

1. Research approach: inductive from case studies

Since the label of open source was formed just before the dawn of the 21st century, the phenomenon of open sourcing while private sourcing (*OSwPS*) has a short history. Theories on the phenomenon of open sourcing have tended to put private sourcing in the background. This research on *OSwPS* is driven by data, towards building theory Eisenhardt and Graeber 2007, as shown in Figure 1.1.

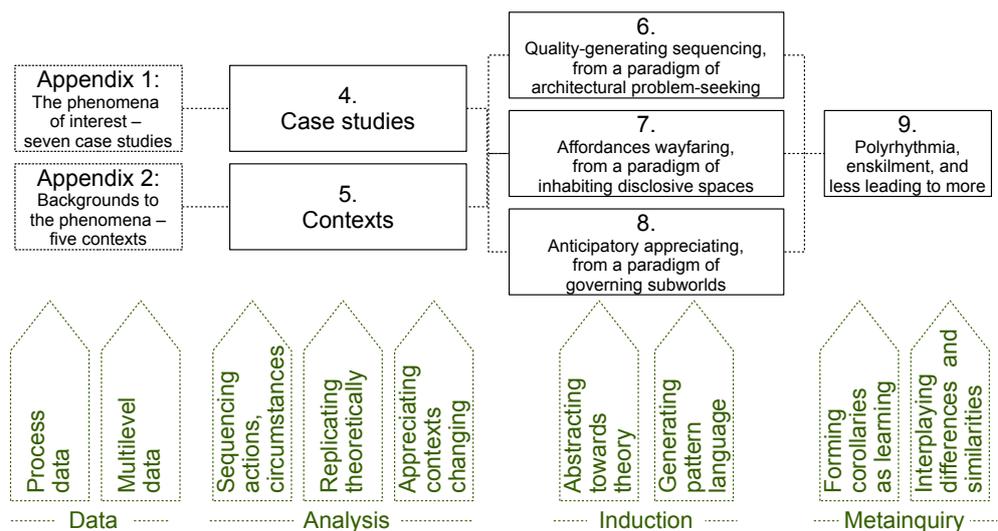


Figure 1.1. Considerations in research approach by chapter

The key considerations in the research approach include:

- *data*, as process data over a decade, viewed as multilevel in Appendices 1 and 2 (section 3.1);
- *analysis*, as sequencing actions, circumstances and outcomes; replicating theoretically; and appreciating contexts changing, across Chapter 4 and 5 (section 3.2);
- *induction*, as abstracting towards descriptive theory, and generating pattern language, in Chapter 6, 7, and 8 (section 3.3); and
- *metainquiry*, as interplaying differences and similarities across descriptive theories, to further building normative theory emphasizing

innovation learning, in Chapter 9 (section 3.4).

The data collected is about and from IBM, benefiting from the public nature of open sourcing information, complemented with insight of the study author as an employee during the period under study. The significant market presence of IBM sweeps in activities of business partners and industry competitors, providing some lessons for others to consider.

1.1 Data: The history of open sourcing while private sourcing is observed as events, activities and choices ordered over time

The data window starts with 2001 as a watershed year, with the January announcement of a plan to invest \$1 billion in Linux over the following three years (IBM 2001, 21). Seven case studies in Appendix 1 (reduced to Chapter 4) provide parallel histories through 2011. Within that time period, the earliest cases progressed from startup to winddown, while the later cases continued beyond the timebox. Five contexts over all of the case studies are described in Appendix 2 (reduced to Chapter 5).

The nature of the datasets has driven their handling as (i) process data, as overlapping time series histories; and (ii) multilevel data involving individuals, workgroups and organizations in both corporate and non-commercial contexts. The considerations for process data and multilevel data are described in the following two subsections.

1.1.1 Process data: Over a decade, ways that open sourcing does and doesn't work with private sourcing were discovered

Insight into evolving phenomena such as organizational learning, innovation and change can be gained through studying data longitudinal over time, rather than as cross-sectional views at a single point in time. When a body of theory is limited and fine-grained qualitative data is available, new theories can be built and later tested with time series and event history methods.

Process data are series of events, activities and choices ordered over time. For researchers, four characteristics that makes them difficult to analyze and manipulate.

First, they deal mainly with sequences of "events": conceptual entities that researchers are less familiar with. Second, they often involve multiple levels and units of analysis whose boundaries are ambiguous. Third, their temporal embeddedness often varies in terms of precision, duration, and relevance. Finally, despite the primary focus on events, process data tend to be eclectic, drawing in phenomena such as changing relationships, thoughts, feelings, and interpretations (Langley 1999).

The data on *OSwPS* in Appendices 1 and 2 follows that list of four

Appendix 1:
Phenomena
of interest –
case studies

Appendix 2:
Contexts
of phenomena –
contexts

Multilevel
data

data

characteristics.

On (i) “data composed of events”, the most interesting aspects are “stories about what happened and who did what when”. Actions and decision made under uncertainty may provide more insight than directions espoused a priori.

On (ii) “data on multiple units and levels of analysis with ambiguous boundaries”, some actions were taken by individuals formally charged with responsibilities and resources, whereas others have relied on guidelines and volunteerism, within and beyond organizational boundaries.

On (iii) “data of variable temporal embeddedness”, the sequences of events occurs at irregular periods, and histories recalled retrospectively tend to focus on significant changes. Individuals and groups learned about *OSwPS* not only within the confines of a single case, but across the organization in complementary or peripheral projects and activities.

On (iv) “data that are eclectic”, variables and events are intertwined. In the cases that have spanned years, the players – both individuals and organizations – have changed. The cases may include unclear and evolving purposes altered from hands-on learning, factors in the larger organization, or unforeseen external pressures.

A study based on longitudinal data in the real world does not have the benefit of data on “the road not taken”. There is, however, a reality of action from conscious choices made by actors immersed in the situation.

1.1.2 Multilevel data: open sourcing while private sourcing coevolved for individuals, teams, corporations and non-profits

The origins for seven case histories on *OSwPS* vary: some started as formally funded corporate initiatives, and others started as voluntary contributions by individuals working on their own time. In an open sourcing style, relatively few people are committed full-time to an exploratory project, yet success is later measured through the widespread adoption of an innovation.

In a multilevel approach, individuals, groups and organizations are embraced simultaneously, with two advantages. Firstly, researchers can avoid significant fallacies associated with single-level empirical research.¹ Secondly, a multilevel perspective opens up new opportunities for theory

¹Fallacies include (i) a *cross-level fallacy* in construct validity, whereby individual-level phenomena give rise to higher-level phenomena; (ii) a *contextual fallacy* in internal validity, finding spurious relationships at lower levels while failing to account for higher-level relationships; (iii) an *ecological fallacy* in external validity, incorrectly assuming that a relationship that exists at a higher level exists in the same way at a lower level; and (iv) an *atomistic fallacy* in external validity, incorrectly assuming that a relationship that exists at a lower level exists in the same way at the higher level (Burton-Jones and Gallivan 2007, 660).

understanding linkages between levels.

In a business working in *OSwPS* at an organizational level, there will be also be individuals and groups concurrently working in open sourcing only (*OSo*) or in private sourcing only (*PSo*). Further, employees can be active with other parties outside of the company participating in *OSo* and *PSo* communities. An approach that includes data from multiple levels encourages an appreciation of the interdependencies that are simultaneously in play in the real world.

With multilevel process data as the source for study, the next section describes the approach for analysis.

1.2 Analysis: In hindsight, processual abstractions of evolutionary stages of open sourcing while private sourcing can be constructed

With the body of *OSwPS* data defined, analysis has been conducted with three considerations: (i) sequencing actions, circumstances and outcomes longitudinally; (ii) replicating theoretically across multiple case histories, and (iii) appreciating the changing context in the background for the duration of the study. In Chapter 4, key *OSwPS* actions and events are charted for each of seven cases over the 10-year period, lined up with some *OSo* and *PSo* periods. While the emphasis is on in-case longitudinal histories, cross-case learning may influence decisions in sister initiatives and teams. Chapter 5 summarizes changing background contexts, as actors learned from larger trends in the world, particularly as open sourcing became more commonplace.

References

- Burton-Jones, Andrew, and Michael J. Gallivan. 2007. "Toward a Deeper Understanding of System Usage in Organizations: A Multilevel Perspective". *MIS Quarterly* 31, no. 4 (): 657–679. ISSN: 02767783, visited on 01/25/2009. <http://misq.org/toward-a-deeper-understanding-of-system-usage-in-organizations-a-multilevel-perspective.html>.
- Eisenhardt, Kathleen M., and Melissa E. Graeber. 2007. "Theory Building from Cases: Opportunities and Challenges". *Academy of Management Journal* 50, no. 1 (): 25–32. ISSN: 0001-4273, visited on 01/15/2009. doi:10.5465/AMJ.2007.24160888. <http://dx.doi.org/10.5465/AMJ.2007.24160888>.
- IBM. 2001. *Annual Report 2001*. Annual Report G507-0501-07. Visited on 05/02/2010. <http://www.ibm.com/annualreport/>.
- Langley, Ann. 1999. "Strategies for Theorizing from Process Data". In this article I describe and compare a number of alternative generic strategies for the analysis of process data, looking at the consequences of these strategies for emerging theories. I evaluate the strengths and weaknesses of the strategies in terms of their capacity to generate theory that is accurate, parsimonious, general, and useful and suggest that method and theory are inextricably intertwined, that multiple strategies are often advisable, and that no analysis strategy will produce theory without an uncodifiable creative leap, however small. Finally, I argue that there is room in the organizational research literature for more openness within the academic community toward a variety of forms of coupling between theory and data. *The Academy of Management Review* 24, no. 4 (): 691–710. ISSN: 03637425, visited on 01/20/2009. doi:10.2307/259349. <http://dx.doi.org/10.2307/259349>.