Archival Research: A Neglected Method in Organization Studies

1   Introduction

Archival research is a much under-rated and under-utilized method of research in management studies. Austin, Scherbaum and Mahlman (2002, as cited in, Aguinis et al., 2009) observe that:

> "In the past 20 years or so, archival designs hover around 10% [of all research designs]. In terms of measurement, classical test theory is clearly dominant and includes 100% of articles until the past decade, when IRT (Item Response Theory) was reported in about 3% of articles."

Note that Aguinis et al (2009) were writing at a time when considerable business history as well as immense quantum of corporate information had already accumulated in the industrialized countries. Seen in that backdrop, the fact that only 10% articles care to make up cursory reference to archival data, appears to be a dismal figure. Taking this premise as our reference point, an attempt is made to survey the literature that has, as yet employed this method and understand the techniques adopted. This paper re-positions the method of archival research as a viable and sophisticated tool for researchers to employ effectively in singular or mixed method studies.

1

2   The Nature and Purpose of Archival Data

At this point, it is imperative, that we define and delimit the boundaries of the construct 'archival data' for the purpose of this discussion. Archival data is that data which is gathered and stored prior to the commencement of the research, intended for later use. Archival data may include census data, court proceedings, patent office records, credit histories, and educational records, among others. Organizational archives may consist of corporate annual reports, personnel files, and survey responses undertaken at different times. Each kind of archive has it's own pros and cons.

The primary and most obvious advantages of archival data are of-course ease of availability and low cost. Organizations typically collect and store copious amounts of data pertaining to employees, customers, suppliers, competitors and others. Human Resource Departments often float questionnaires for feedback from time to time on various issues of importance for the organizations. Research and development departments keep records of their own progress, patents etc. while legal departments tend to store proof of all types of legal transactions undertaken by the company or company's agents. MacCallum (1998, as cited in Payne, Finch, and Tremble, 2003) says that it makes eminent sense to use these high-quality data sets as "a way to test new ideas on existing data, often with large samples, while saving time and other resources". The cost factor is perhaps the single largest determinant in favour of archival research. There is sometimes a fee required to obtain access to such datasets, but this is almost always a tiny proportion of what it would cost to conduct an original study (Cheng and Phillips,

2014). This cost saving is not only valid from the point of view of the researcher who saves on substantial travel and implementation costs but also from the perspective of the management, because there is a visible utilization of the initial investment and consequent satisfaction. A third advantage of archival data is the default largesse and richness of the datasets. Fourth, it is particularly suited for certain research questions or types of variables that either can't be manipulated ethically or can't be studied on account of logistical constraints, one example being workplace accidents. Fifth, archival data may be easily used to replicate research conclusions, as such they enhance the generalizability of theories. Finally archival data is frequently of a longitudinal nature, something that is very painstaking to collect when using primary research methods again owing to time and cost constraints of the scholar. Increasingly, as longitudinal studies assume more importance in Organization studies; in the determination of antecedents, correlations, and consequences of psychological constructs; archival data is likely to gain in attention and credence.

Within the social sciences, archival analysis techniques have long been considered a non-invasive technique of gathering requisite data. Zwijze and Jong (2005) document several methods in this ambit. Citation analysis is one of the most popular ways of tracing citation networks among ideas, scholars and research programs, which we take up again later. One very commonly available and vital form of archival data are communication records such: internal mail, phone records, and e-mail exchanges. Because the variables in such data like initiators, recipients, and dates are usually available and can be coded, these are very amenable to quantitative methods of analysis (as also qualitative). Content analysis is a method that is frequently used by communication researchers. At a more complex level, these communication

3

records may also be used to uncover underlying groups or power bases within the organization. Hence depending on the granularity level of the archival data available, varying degrees of sophistication in research questions is possible. As Layder (as cited in Welch, 2000, p.172)) argues, it also has a place in 'contemporary, theory-generating research'. It can be argued that the role which archival data may play in such contemporary-oriented research is threefold : (1) archival data can be used to add 'empirical depth' to a project by generating new data and enabling verification of existing data from other sources; (2) archival data are particularly suited to generating developmental explanations, in other words, explaining processes of change and evolution; (3) archival data can be used to challenge existing theories and build new theoretical models.

It is also prudent to look up the disadvantages associated with archival data, while determining it's feasibility. First, the data may have been collected simply for documentation purposes and mostly nonscientific reasons; as such it is likely that it may not have all the variables that are of interest. Second, quality may be compromised by missing data points, non representativeness, and/or errors. Third, archival data is highly susceptible to experimenter bias i.e. researcher may be tempted to examine the data and accordingly formulate convenient hypotheses. Fourth, restricted access documents can be much more difficult to obtain (Hughes and Griffiths, 1999). To protect the confidentiality of respondents, publicly available datasets usually delete identifying variables about respondents, variables that may be important in the intended analysis such as zip codes, the names of the primary sampling units, and the race, ethnicity, and specific age of respondents. This can create residual confounding when the omitted variables are crucial covariates to control for in the secondary analysis (Cheng and Phillips, 2014). Finally,

4

sometimes, this data may not be amenable to analysis i.e. it may have reliability and validity issues.

Charlotte (2016) argues that we should treat documents as more complex forms of data. This in turn necessitates that we be mindful of potential access issues; reflexively consider the emotional risk of using documents in research; take into cognizance the relevance of ethical issues, such as researcher safety, and reflect on the researchers' emotional engagement with actors discussed in these documents. These considerations will lead to a more nuanced engagement and understanding of the potential social and emotional impact of working with sensitive documents. Harris, (2002) eclipses the gist of the debate with this observation about the 'nature' of archives:

> " [...] the idea is that archives, mirror-like, reflect reality [...] Of course, the assumption that there is "a reality" capable of reflection in records is debatable from a number of perspectives [...] First, even if there is "a reality," ultimately it is unknowable. The event, the process, the origin, in its uniqueness,is irrecoverable [...] Secondly, while it is self-evident that the record is a product of process, it must be acknowledged that process itself is shaped fundamentally by the act of recording. And thirdly, if archival records reflect reality, they do so complicitly, and in a deeply fractured and shifting way. They do not act by themselves. They act through many conduits – the people who created them, the functionaries who managed them, the archivists who selected them for preservation and make them available for use, and the researchers who use them in constructing accounts of the past."

5

3   Method

3.1   Stage 1: Benchmarking Articles

While there is prescribed methodology and accessible guidance to search and acquire literature in areas where a considerable body of knowledge (read empirical and theoretical papers) has accumulated, known as the 'Systematic Review protocol' (Cook, Mulrow and Haynes, 1997); how to search literature for a topic that has seen very little to almost non-extant research and development? This question is a double edged sword. While the topic itself looks to be an appealing research gap owing to it having been scarcely explored; on the other hand, backing it up with 'evidence' and presenting it to one's peers as a 'Justified True Belief' (Lehrer and Paxton, 1969) becomes that much harder. To our consternation, we found that this search for evidence became a mini project in itself. Since, a traditional mode of literature search in EBSCO or Google databases was unamenable to our question of formulating an archival methodology protocol; we therefore improvised an 'iterative and consultative' search strategy. To formulate this strategy, we asked the following questions:

1. Which are the emerging domains or sub-disciplines within organization studies or close to it that might benefit from a protocol for archival research?

2. Do researchers in these domains already have a protocol handy, do they improvise or are they looking for one or are unsure? Do they have method papers relating to this problem for guidance? Can we acquire those papers from them?

3. Are these method papers: (a) consolidatable in the sense that do they have an emerging unifying, common technicalities, or are they (b) idiosyncratic in that they are contextually embedded in the discipline to which the particular research questions bear fidelity?

6

4. If the answer to the above is 'b', how do we formulate the guidelines? Do we take a distributed approach or do we take a unified?

Our questions yielded the following set of consequential recourses to us. First, Our survey of literature, based on thematic analysis revealed, that there are three emerging sub-disciplines 'within' organization studies (i.e. Business History, Social networks and management Education) and one major discipline (i.e. strategy management) closely aligned to it that are currently seeing an influx of archives based research studies. Additionally, another area, accounts and audit, which was not close in epistemological origins to organization studies was also witness to this phenomenon because of maturation of the field and an interesting cross-fertilisation with eclecticism in the form of Critical theory paradigm (c.f. and see mission statements of Journals: *International Journal of Critical Accounting*, *Critical Perspectives On Accounting* and *Accounting and Organization*). Second, we floated the question # 2 amongst peers from different domains such as Marketing, Information Systems, Operations, Strategy, Finance and Economics, in our institution and inquired about the prevalence of usage of this method as well as accessibility of protocol. Only Information Systems, Strategy and Finance researchers came back with affirmative answers about the first part of the query; also they provided us with the papers that we subsequently used as benchmark articles for this paper in the section: 'Multi-Disciplinary Undertakings with Archives. All of them concurred on the point that there was as yet no best-practice recommendation on methodology available. An interesting finding of our peers based focus group discussions was that Knowledge Management (KM) emerged as a contemporary development on which the perceptions of potential for this method converged amongst both organisation studies and information science researchers. Note that our strategy pays us off twice

7

over. Not only does it, by default, eradicate the 'file drawer' problem; but also facilitates an early

stage peer validation of the search and benchmarking process.

3.2 Stage 2: Content Analysis and Synthesis

Having benchmarked our core articles of interest, we turn to the task of synthesizing the

techniques employed in them and devising our protocol. This also brings us to addressing

Questions (3) and (4). Our content analysis revealed that in terms of instruments and/or steps,

there were some common underlying principles that could be translatable in all the five domains

of interest. These, we compiled under: 'nature, purpose, modes, toolkits and varieties' of archival

research. We also dealt with intricacies such as minutely examining the fieldwork steps, proxies

generation and other related processes of triangulation in this part. In the second part of the

distributed approach that we eventually took, we listed our findings according to each discipline

header. As can be expected, there were eccentricities in each discipline in respect of aspects such

as data sources, among others. We also reckoned that this approach would serve to open up our

readers' minds towards the possibilities and varieties of future research questions pertaining to

each research domain covered. Finally, coming into the more contemporary concern of KM; we

have attempted a survey of KM's technological applications, issues and challenges and what

implications they hold for management researchers.

[Insert Figure 1 ('Method Flow') about here]

8

- -

4   Toolkit of the Archival Researcher

The literature delineating precisely the method to do archival research is very sparse; the general consensus being that it is open to intuition and not complex enough to require documentation. However, we could come across a few well-articulated methodological papers which we cover under three subthemes, viz. (1) Archival fieldwork (2) Archival proxies and (3) Of Mixed methods, triangulation, reliability and validity.

4.1   Archival fieldwork

Welch (2000) underscores the point that it is extremely important that archival researchers follow a pre-defined protocol, which he outlines as a 5-step approach : (1) Discovery: this involves tracing the whereabouts of a collection. Often it happens that the company or present employees are not even aware if a certain record exists. In that case, one may need to track down former employees and elicit information. (2) Access: Sometimes a company may simply deny the existence of information because of perceptions of the nature of sensitivity of the said record. Again, then, one may need a former employee on one's side as a 'sponsor' for the project. (3) Assessment: Once access is secured, the records need to be examined properly to ascertain if they are usable in one's research. For this, we need to look into 4 counts of quality: authenticity, credibility, representativeness and meaning. (4) Sifting: This pertains to reducing the volume of information to a manageable quantum and involves 2 steps: Data reduction (sequencing techniques) and reconstruction (patterning techniques). (5) Cross-checking: This is the final step

9

and refers to triangulation both data-wise and methodologically. Bridges et al., (1993) notes that developing a knack for understanding some of the basic archival principles may come in handy while locating evidence. On the other hand, with specific reference to electronic archives, other scholars have demonstrated the value of 'indexing' and other associated techniques (Feeney, 1999; Yakel, 2003).

4.2   Archival proxies

Ketchen, Jr., Ireland and Baker (2012) define an archival proxy as "a quantitative measure that is used to represent a theoretical construct that is relevant to the design and completion of a research study". Scholars have demonstrated that archival data may be used to develop proxies to represent psychological constructs to be used in organizational behaviour research. This concept of 'Surrogate Measures' from archival data for (behavioural) constructs of interest or 'Archival proxies' has been dealt in substantial detail by Payne et al (2003). To illustrate the concept as well as the application, they describe the dilemmas they faced from their own research:

> "We wished to examine the longitudinal effects of organizational commitment on turnover intentions and actual turnover behaviour within the army. Few studies have examined these relationships beyond the first year of employment; hence, there was a need to examine this relationship over a longer time interval. Second, we learned of a data archive containing survey data on army officers' career decision making over the course of 12 years, which appeared potentially suitable for longitudinal research on the development of organizational commitment—hence, the opportunity to test related research questions. However, the archive contained questionnaire items that sought officers' views on a number of practical issues, not previously validated measures of psychological constructs. Fortunately, some of the

10

questionnaire items appeared to measure organizational commitment and could be used to derive surrogate measures of this construct. We also had an opportunity to collect a validation sample allowing us to examine the validity of the surrogate measure and construct equivalence."

There are five ways to develop archival proxies (Payne et al., 2003), such as by (a) aggregation of archival variables judged by subject matter experts (SMEs) to be consistent with the concerned psychological construct (b) aggregation of those items that have the most robust psychometric properties, (c) Formulation of an abbreviated scale, (d) Usage of a single-item measure, and (e) Usage of a demographic variable or a mix of background variables. Another approach pertaining to the development of surrogate scale is especially creative and hence interesting. It is done by the common method of exploratory factor analysis followed by confirmatory factor analysis i.e. regressing the (archival) items against a previously validated scale and assessing their relative robustness based on the factor loadings.

Once the 'surrogate measure' is developed, it is important to check whether it passes the construct validity test i.e. whether it captures the construct it is supposed to. Both construct under-representation i.e. not capturing all the attributes of the construct in the surrogate and over-representation i.e. crafting too broad a surrogate variable to capture meaningful impact are to be avoided. The literature recommends that applying the construct equivalence approach may effectively address this issue. It is a good idea here to take into cognizance some of Boyd, Dess, and Rasheed's (1993) advice as to the precautions requisite when differentiating between archival and perceptual proxies.

11

4.3   Of Mixed methods, Triangulation, Reliability and Validity

"Mixed method studies have the potential to address research topics that other methodologies cannot attend to on their own" (Williams and Shepherd, 2015). Scholars have tended to enrich their studies with archival research implying that they have used it in conjunction with other methods. As an illustration, let us consider Welch's (2000) paper on business networks. Because he has used archival data as a compliment to the interviews of the network participants, he has dealt at length on how they complement each other. Thus he observes that

> "Archival sources generally have three advantages over interview data: they are more detailed, less obtrusive, and less contingent. To turn to the first characteristic, archival material, unlike interviews which rely heavily on faulty human memory (Huber and Power, 1985) often provide a very precise and immediate record of the interfirm interactions being studied. The second advantage is that archival research is unobtrusive and non-reactive, especially when compared to the social interaction which is a necessary part of conducting an interview. The third advantage of archival research is that it is often less contingent than interviews, which can be affected by a myriad of extrinsic conditions: time constraints, interruptions, the presence of third parties monitoring the discussion, sudden crises facing the company which distract the interviewee, and so on."

It is important and useful to conduct an analysis such as the one quoted above, to justify any kind of methodological triangulation undertaken in a particular study.

Treading on similar lines, Lievrouw, Rogers, Lowe, and Nadel (1987) made an attempt to pinpoint the contribution made by each method to their understanding of networks among biomedical scientists. The archival data used in this case were grant proposals and co-citation

12

analyses, which complimented a set of sociometric questionnaire administered to the respondents as well as individual interviews pertaining to thenature of their contacts. While the sociometric questions overlapped in their utility with the archival analysis in that they were able to identify broad clusters of scientists, however, when it came to the exact nature of substantive collaborations based on co-occurrence of ideas, it was the archival data that revealed the underlying and more intricate patterns.

As far as reliability of archival records is concerned, as long as the research pertains to core network variables (eg: patterns, groups, and roles) collected tacitly, they are not biased by employees' comprehension of questions, their memories of events, or attitude of social desirability of answers (Payne et al., 2003). Hence, whether the data is reactive will largely be a function of what policy the organization follows in informing employees about the Research, this is because employees' behaviour in mail, phone calls, e-mail and questionnaires may change if they are aware of being studied. Again with respect to feasibility of the method, and this is connected to the previous discussion as well, the primary concern is how to deal with privacy issues, especially when the content of messages or biographic data of employees is also analyzed. On the technical side, database management systems of the organization must be kept up-to-date and free from data corruption caused either by human or system errors.

We have already come across the issue of data and methodological triangulation in the fifth step of Welch's protocol under the heading 'Archival Fieldwork'. However, we would like to point out at this point, that we have strengthened our claim by a survey of Knowledge management (KM) articles in table 1 in the last section of this paper where we document the different

13

combinations of methods and data that archival research has been used in, as a reliability and validity enhancing component.

5 Multi-disciplinary Undertakings with Archives

While archival data time and again has been used across the entire span of management research, we would like to highlight five studies that stand out as exemplars. The five studies pertain to five distinct areas, viz. (1) Business history, (2) Social networks and social capital, (3) Management Education and Archival Paradigms, (4) Accounting studies and (5) Strategic management. Additional papers of relevance are noted, where applicable.

5.1 Business history

Business history is of-course the most obvious choice of discipline for an archival researcher and in recent times, this field has witnessed a resurgence of interest in academic circles. Armstrong's (1991) paper is a testimonial to that resurgence as he lists out in detail all the numerous record centres that a keen student of business history may visit in London to really know about British Industrialization. The author makes his case for the article by observing that locating records related to a particular company or group of firms in a specific record office can be a rather tedious and time consuming process. It is therefore handy to have a ready compilation that can be accessed by all business historians anywhere as to where they should go depending on their specific interest area/time/research question. A short recap of Armstrongs' descriptions will make a good illustration for future such compilations in other geographies. Thus the prominent national archives are: (1) The Public Record Office – it is the single largest collection of data

14

about "dead companies" i.e. companies gone bankrupt; (2) The-Guildhall Library – is the largest Stock Exchange data repository comprising of Annual Reports, prospectuses etc. from 1875 onwards.(3) Companies House – this is the office of the Registrar of Companies, established by the Companies Act of 1844 and comprises all records lodged by currently-live companies throughout their existence for over 130 years; (4) The Colindale Newspaper Library – laos called as the British Library, Newspaper Division, but being located separately in north London is called so. It maintains copies of all British daily national newspapers, local dailies and weekly newspapers, yellow pages and liberal press. An additional item of interest is English newspapers from the erstwhile British colonies; and finally (5) British Library - Science Reference Division – housing a comprehensive coverage of all patents and trademarks. In addition to the five, highly visible, national repositories, Armstrong also points out several eccentric ones such as the Country record offices, libraries maintained by Professional associations, trade unions, NGOs and others, each of which have their unique and specialized collections.

This illustration serves to highlight another important, related point: that there is a lot of untapped 'Grey Literature' lying around in many disciplines (Alberani, et al, 1990; Banks, 2006). Scouting for, locating, identifying, cataloguing and harvesting this information is something which again underscores the importance of having a good fieldwork strategy. As business history makes inroads into diverse specialisations from information systems (Buckland and Liu, 1998) to international business (Jones and Khanna, 2006; Buckley, 2009) and onwards to comparative international accounting (Carmona, 2004), so do archives. We reckon that what business history has done to organisation studies in terms of deepening the body of knowledge; archival research will do to historiography (Jacques, 2006). By delineating clearly the multiple

15

designs in which organisational researchers can navigate history, Rowlinson, Hassard, and Decker, (2014) also give a fair idea about what kind of archival data to look for, given a particular level and granularity of historical analysis.

5.2   Social networks and Social Capital

We take Welch's (2000) paper as a representative study in this research domain. This project, a longitudinal case study of the network relationships maintained by a sugar exporting company in the period 1950-1985, involved interviews with key participants in an industrial network, as well as the extensive use of archives in eight documentary collections located in both Australia and Britain. These collections comprised the records of a variety of organizations - an overseas customer, four government authorities, an industry association and an interest group - as well as those of the focal company. The research strategy used, therefore required familiarization with a range of archival settings, as well as a comparison between data gathered through interviews and archives. An additional cite may be Williams and Shepherd (2015) who push the boundaries of social network research by amalgamating methods as diverse as organizational histories, the Gioia et al., (2012) methodology of manual content analysis and inductive data structuring driven by extant literature. What they achieved were detailed social ties maps for 143 new ventures explicating how and between whom brokerage structures emerged to make the businesses viable. The crucial takeaway in organisational level network studies is that once we agree that each 'Ego' or focal organisation's actions are heavily influenced by the structure that it is embedded in (Granovetter, 1985), then acquiring as much information as possible about related parties, stakeholders or 'alter egos' i.e. here, other organisations; becomes a vital

16

requirement. Peters (2010) even helps to establish the dormant networks of the 'troopergate' scandal through his own version of a similar methodological combination.

5.3   Management Education and Archival Paradigms

5.4   Human societies have their own evidence and memory paradigms which guide them to decide what is to be remembered and what is to be forgotten (Mckemmish et al., 2005). Mckemmish et al (2005)   refer to these as 'Archival Paradigms' which give broad guidelines as to how archival knowledge is to be defined, what forms the archives should take, how they are to be described and indexed, and who would have ownership, custodial and access rights to them. These in turn determine the archival properties of reliability, authenticity, and trustworthiness. The concept of the 'Communities of Memory' is thus germane to archival science as well as concurrently and consequentially to the broad education infrastructure extant in society as well. The common past of the society or community is embedded in the memory texts [in any form, written, oral, as well as physical] and thus survives time. In other words, how much a community takes care of it's records is an indicator of how much it values it's past and accumulated education. Major research and development projects in the West have repeatedly underscored the philosophy of these initiatives. In the non-western world, there has been a call for development of indigenous archival practices i.e. local translation of the global archival paradigms. For instance, in South Africa, archivists are exploring "the archive outside the archival inheritance of colonialism, and later, apartheid" – the oral record, literature, landscape, songs, dance, ritual, art, artefacts and so on.

17

This will redraw the boundaries and relationships between power and freedom. In another instance, an Indigenous Australian community's meteorological classification of climate as hot or cold weather patterns, identified in terms of the growth of plants or animal behaviour, was harvested in the form of a calendar to synchronise with Western notion of seasons. It opened up many insights into the flora and fauna behaviour in that way. Essentially, this (education from collective memories) is one area which is largely untapped in archival studies. Organizations being social systems, it would be interesting to explore what the collective memories of particular organizations may entail for the education of the larger society. Becker (as cited in Decker, 2013) avers that "archives are the memory of an organization". We can deduce from the work of multiple scholars (Brothman, 2001; Decker, 2013; Decker, 2014) that like human memory, keeping track of organizational memory is also therefore a tricky path. Accounts and Audits

In our fourth domain of archival interest, i.e. accounts and audits, we take Maines and Wahlen's (2006) paper as an illustration. The authors submit that accuracy and reliability of accounting information is an area which is very difficult to establish. It is however extremely important to do so for a number of stakeholders: standard setters so that they can determine requisite financial reporting standards, account preparers, auditors, and financial statement users (such investors and creditors). In the context of accounting data, reliability itself has three characteristics - representational faithfulness, verifiability, and neutrality. But however, this reliability is again a matter of degree. As a first test, accounting data needs to match the underlying economic constructs that determine the future cash flows to the firm. As an example, the authors cite Chandar and Bricker's paper (2002) which demonstrates that returns to market-wide portfolios

18

(i.e., the S&P 500 Index and the Russell 2000 Index) act as a useful proxy to gauge the reliability (e.g., overstatement or understatement) of the fair value estimates of gains and losses projected by a sample of closed-ended mutual funds'. For accounting standard setters, the reliability of archival accounting research assumes a whole lot more of importance. Bodies such as the Securities and Exchange Commission regularly report on firms that violate GAAP (Generally Accepted Accounting Principles) or refurbish their accounts. This is extremely vital and usable archival information. Studies and reports even examine the components of earnings (such as revenues), which may be fudged most, leading to the maximum negative market reactions, the greatest likelihood of enforcement actions and litigation, and largest settlement costs. These areas thus emerge as red flags for standard setters.

Maines and Wahlen, (2006)go on to outline specific techniques to assess the reliability of accounting data, viz: (1) comparison of the accounting numbers to economic benchmarks, (2) examination of restatements of accounting data, (3) comparison of accounting projections against future cash flow realizations, (4) investigating external incentives to see if they link back to reported accounting numbers and (5) gauging reliability from the patterns of usage of the accounting data in the capital markets and other contexts. See DeFond and Zhang, (2014) for a comprehensive overview of archival research techniques in auditing research.

5.5 *Strategic management*

In our final area of interest, we take up Ketchen Jr. et al's (2012) study for illustrative purposes. The authors assert that the inherent nature of constructs investigated in strategic management research require a supra-observational conceptualization, as such the usage of archival proxies is

19

essential and pervasive. Taking the example of three archival proxies in particular: (1) research and development intensity, (2) patent counts, and (3) patent citations; they demonstrate that multiple authors have used them to represent multiple strategic and organizational level concepts in the literature. Herein, according to them also lies the problem. For instance, R&D intensity has been variously used to represent technological knowledge (Choi and Wang, 2009), Differentiation strategy (Graffin et al., 2011) and Strategic investments (David et al., 2006), among other things. Let us assume now that a scholar is interested in investigating the link between strategic investments and market share and for that she decides to use R&D intensity as a proxy for strategic investments. However, on the other hand, there surfaces an empirical study that conclusively shows that R&D intensity is related to market share. In this case, then whether it will represent strategic investments or any of the other afore-mentioned alternatives, will become a moot point. The authors point out that herein becomes relevant, the role of 'gatekeepers' such as editors or reviewers in the censoring of arbitrary practices and of clumsy usages of precious archival data. We therefore feel that this last domain presents an altogether different conundrum to management researchers. While it is already an established fact in this case, that archives and archival data is a mainstay for empirical endeavours here; there is also a case for caution in that there is a felt need for suitable norms and checks and balances for the effective operationalisation, capturing, reporting and theoretical usage of these resources.

Additionally, Mintzberg and McHugh (1985) use archival research, supplemented by interviews, to trace the development of organizational strategies over time. Taking a similarly processual view, Pettigrew (1985), in a well-known study, assesses strategic change in the British company ICI over a 23-year period by combining archival research with interview and observational data.

20

Audia, Locke, and Smith's, (2000) paper triangulates the archival study with a laboratory based one to delineate the paradoxes inherent in persisting with a strategy in the face of massive environmental shocks.

6 Knowledge Management's Technological Applications

We have already implicated 'memory', human and organizational, in the information management applications being developed by and for organisations. To say that the exponential increase in sophistication of computing has made it much easier and sustainable to store data in digital drives rather than physical tomes, is redundant. At a deeper level, what has happened is that, the discipline of Knowledge Management (KM) has emerged in a big way not just as a practice in corporate culture and corporate knowledge conservation, but also as a burgeoning research discourse..Levinthal and March, (1993) explain this intertwined progression in KM thought and practice thus:

> "The (re)discovery of learning has been stimulated by the current interest among students of strategic management in organizational capabilities and knowledge [...] Because strategic anticipatory rationality seems to demand both greater cognitive and calculative capabilities and more consistency and stability in preferences than can be reliably assumed, considerable effort has been directed to improving the informational and analytical basis for organizational action and to developing consistent, stable organizational objectives. Modern decision-oriented information systems and procedures for defining (or negotiating) goals reflect this spirit"

21

As is evident from the quote, the very structural properties of this body of research, by default, make it amenable to archival improvisations.

6.1 Sampling of Empirical archival research articles

In continuation of our discussion on 'Multi-disciplinary Undertakings' and also to provide an overview of current knowledge management research, we tabulate a selective sample of empirical papers from the domain to illustrate the diversity of usages and research methods possible with archival data, in Table 1 below:

- - - - - - - - - - - - - -- -

[Insert Table 1 about here]

- -

The selection of compatible methods, as illustrated above, is also heavily influenced by the increasing complexity and diversity of knowledge management related inquiries. We explore some such questions in the next section.

22

6.2 Evolving KM Questions

One of the exemplars of research into the substantial changes that modern work environments and employee behaviour have undergone in response to the digital onslaught is Ruhleder's (2000) paper on Distributed environments. Ruhleder ethnographically studied two separate contexts, one, a distributed corporation: The Holding Company (THC) and another, an online master's degree program offered through the University of Illinois at Urbana– Champaign and delivered as 'LEEP'. Researchers studying distributed settings are typically focused on three sets of issues: (1) the network practices emerge out of participation and cooperation in these virtual and particularly hybrid environments; (2) the organizational policies and cultural barriers etc. which accompany these technology changes and (3) the re-defining of roles, relationships and boundaries. Some other names for the technology bundles that get utilized in the two contexts are: Human-Computer Interaction, Computer-Mediated Communication, Management Information Systems (MIS), and Social Informatics.

In the first case i.e. THC, Ruhleder's team used their own video equipment i.e. independent of the VLAN setup in which cross-location teams were conducting a meeting. This allowed them to record what was seen or heard locally as side-talks and gestures as well across the link. In the LEEP case, they would log in to the chatroom associated with the virtual classroom. That way, they could follow the entire discussion among students while the lecture was delivered parallely.In the Post Hoc Archival Analysis, the authors collected and analyzed all the audio-chat events and the class conference postings that are archived in the university website. This is thus an example of Virtual Ethnography at it's best. The authors note that several corporations have tapped into the latter mode of enhancement of organizational learning and knowledge resources

23

i.e. LEEP. Examples are Xerox, Carnegie Mellon, Bell Labs, the University of Michigan, the University of Toronto, and others.

In a comprehensive overview of the status of Archival science, Couture and Ducharme (1998) report that archivists have their work cut out for them. David Gracy (as cited in Couture and Ducharme, 1998) sums up the imperative for the archivist and the process of archiving in the following five conditions:

> "It has to move beyond simple narratives and apply an appropriate methodology such as comparative, statistical, qualitative, or historical analysis; the archival nature of information must be a priority research field; it must be supported by the information sciences, especially when electronic documents are involved; it must be international; and systematic efforts must be made in order to find the necessary funds to finance the research."

Since the function of records management links to the broader aspect of organizational administration as well, therefore scholars have listed four possible fields of study and development in the management of archives: the first pertains to the theme of corporate culture; how may for instance archivists ensure that employees consult records more often? The second issue may be of the systematic study of the organizational effectiveness of archival services. The third question is whether or not there should be regular archival audits by management and last whether given a particular organization's present structure and future potential, there should also be review of rebooting of the archival function in the company. Of course, with the advent of digitization, all of the above questions have assumed even more importance for archival management than ever before.

24

On the academic front, as social science research has proliferated; the academic world realizes that there is a need to obtain and conserve the datasets from individual researchers so that they can be made securely available to the larger scientific community without the danger of information loss or corruption. This danger arises because "Data sometimes exist on individual researchers' Web sites, without professional backups, off-site replication, plans for format conversion and migration, or professional cataloguing." King (2007) makes a strong case for the creation and maintenance of a centralized repository or archive housing such as 'replication datasets' when he says that "Articles with accessible data are cited twice as often as otherwise equivalent articles that do not provide data access. Journal editors who require authors to make data available are also acting rationally in advancing the success of their publications: Articles in journals with replication policies that make data available are cited thrice as frequently as otherwise equivalent articles without accessible data". King thus proposes a cutting edge solution in that a 'Dataverse' be formed not unlike a Google search engine and Google server combined. What would be novel however, about this Dataverse, compared to a normal digital archive; is that on searching online, it would return the required data from the dataverses or websites of individual authors themselves replicating the hosts' look and structure but the actual physical location of the data would be offsite at a neutral, collectively operated and incorruptible server.

Macqueen and Milstein (1999) bring out a startling paradox from the woodworks when they say that the explosion in qualitative research has curiously necessitated a grip over more positivistic methods and tools such as information system understanding, data management and data analysis. Delineating the peculiar challenges that qualitative data sources such as video, audio

25

and multimedia present as opposed to quantitative sources such as spreadsheets or reports; he asserts that more specific and robust KM architecture is required if we want to effectively handle the storage, usage and retrieval of this knowledge. They thus recommend a four-component design, consisting of "(1) characteristics of the sources where information is sought, (2) primary information or objects collected from the sources, (3) secondary information or objects created to aid in the interpretation of primary objects, and (4) characteristics of the coders who construct the secondary objects".

Koretz and Lee (1998) in an interesting observation, peculiar to the multi-billion dollar drug and pharma industry, note that installing and implementing KM solutions in drug companies come with their own challenges. The MEDLINE database that all companies rely on and which is run by the National Library of Medicine (a US non-profit) covers some 3,900 journals and nine million records of clinical trials. This data is painstakingly indexed and catalogued so that users can easily search for particulars tests of their interest using the MESH terms. Subsequent to the search, companies use expensive softwares such as (1) 'Computer-aided trial design' (CATD) - to evaluate the range of possible outcomes for a proposed clinical trial of the candidate drug; (2) Computer-aided program design (CAPD) - to assess alternative clinical development programs and finally (3) Clinical data management systems (CDMS) – to efficiently store and retrieve the data gathered in clinical trials, thus completing the cycle.

Koretz and Lee note that while the database itself, now known as PubMed is freely available, the unique challenge in drug development KM arises however not from storage or retrieval functions but from the capture function. Scholars assert that in an effort to make the voluminous data more

26

accessible to retrieval processes, the current MESH architecture which follows a domain knowledge design has made the capturing and intricate indexing processes extremely expensive. Hence, they deduce that ontological clarity in a particular knowledge domain becomes a crucial deciding factor in designing a cost-effective KM system. What is today unique for drug research, will become more common tomorrow for social science research as well, as knowledge accumulates. In that possibibility, the idea resonates very closely with that of the 'Archival Commons' as envisioned by (Anderson and Allen, 2009). Having said that, we must however bear caution that for all domains of study, digital KM applications can support us only so much as traditional research rigor feeds into it.

6.3   *The Advent of Scientometrics*

The offshoot of the digitization of knowledge produced has also meant that all data pertaining to the dissemination and subsequent use of that knowledge can also be traced and preserved. We know this as 'metadata' and some manifestations of these are the 'doi' links of journal articles, affiliation and other details of authors and co-authors, citation and co-citations mentioned in the articles and so on. Sociologists of science already agree that science is a co-creative process (Latour and Woolgar, 2013) and that no knowledge happens in a vacuum. What this and the emergence of meta-data implies is that it is possible for us now to capture the trajectories of how knowledge flowing is in different disciplines, the waxing and waning of it as also the accumulation and diversification of different 'schools' of thought, more rigorously with proper 'metrics' to capture the data. For instance, we can track the influence of a 'school' of thought across time and geography by plotting the social networks among co-citing authors. The

27

techniques and phenomena we have just described are what comprise the 'science of science' or Scientometrics. Digital archival data essentially has spurred an entire disciplinary program of research. Based on available evidence, we therefore offer a conjecture that probably the meta-discipline that will most benefit from the digital revolution or 'data deluge' as scholars (Meyer and Schroeder, 2009; Buneman, 2005) call it, will be that of Scientometrics.

7   Reflections

We have tried to establish through this article a favorable disposition among researchers towards archival research as a complimentary tool. Specifically, we expect that we have achieved the following four objectives with this paper. First, we have attempted to develop a steps based protocol for archival research, beginning with the process of fieldwork that involves scouting for data; then moving on to the processing stage where data cleaning, proxy generation etc are undertaken and finally culminating in the process of triangulation of the research endeavour. Second, we have attempted to critically compile a trans-disciplinary corpus of best practices in archival methods. This corpus itself is discipline-agnostic and may be used by future researchers as a ready reckoner for further improvisations in this methodology. Third, we have discussed the ability of this method to handle with elan, both qualitative and quantitative data. Finally, we have attempted to open our readers' minds to the possibilities in the emerging domains of knowledge management and scientometrics and the considerations involved. Methodological richness leads to substantive granularity. As such, we argue that archival methods contribute to the robustness, contextuality and holism of any research endeavour, more so in the study of business and organizations.

28

8   References

Aguinis, H., Pierce, C.A., Bosco, F.A. and Muslin, I.S. (2009), "First Decade of Organizational Research Methods", Organizational Research Methods, Vol. 12 No. 1, pp. 69–112.

Alberani, V., Pietrangeli, P. D. C., & Mazza, A. M. (1990). The use of grey literature in health sciences: a preliminary survey. Bulletin of the Medical Library Association, 78(4), 358.

Anderson, S.R. and Allen, R.B. (2009), "Envisioning the Archival Commons", The American Archivist, Vol. 72 No. 2, pp. 383–400.

Aribi, A. and Dupouët, O. (2015), "The role of organizational and social capital in the firm ' s absorptive capacity", Journal of Knowledge Management, Vol. 19 No. 5, available at:https://doi.org/10.1108/JKM-05-2015-0169.

Austin, J. T., Scherbaum, C. A., & Mahlman, R. A. (2002). *Handbook of research methods in industrial and organizational psychology*. (S. G. Rogelberg, Ed.). Oxford, UK: Blackwell.

Armstrong, J. (1991), "An Introduction to Archival Research in Business History", Business History, Vol. 33 No. October, available at:https://doi.org/10.1080/00076799100000002.

Audia, P.G., Locke, E.A. and Smith, K.G. (2000), "The paradox of success: An archival and a laboratory study of strategic persistence following radical environmental change", Academy of Management Journal, Vol. 43 No. 5, pp. 837–853.

Austin, J.T., Scherbaum, C.A. and Mahlman, R.A. (2002), Handbook of Research Methods in Industrial and Organizational Psychology, edited by Rogelberg, S.G., Oxford, UK: Blackwell.

Banks, M. A. (2006). Towards a continuum of scholarship: the eventual collapse of the distinction between grey and non-grey literature. Publishing Research Quarterly, 22(1), 4-11.

29

Boyd, B.K., Dess, G.G. and Rasheed, A.M.A. (1993), "Divergence between Archival and Perceptual Measures of the Environment: Causes and Consequences", Academy of Management Review, Vol. 18 No. 2, pp. 204–226.

Bridges, E., Hunter, G.S., Miller, P.P., Thelen, D. and Weinberg, G. (1993), "Toward Better Documenting and Interpreting of the Past: What History Graduate Programs in the Twenty-First Century Should Teach about Archival Practices", The American Archivist, Vol. 56 No. 4, pp. 730–749.

Brothman, B. (2001), "The past that archives keep: memory, history, and the preservation of archival records", Archivaria, Vol. 1 No. 51.

Buckland, M. and Liu, Z. (1998), "History of information science", in Hahn, T.B. and Buckland, M. (Eds.), Historical Studies in Information Science, pp. 272–295.

Buckley, P.J. (2009), "Business history and international business", Business History, Vol. 51 No. 3, pp. 307–333.

Buneman, P. (2005), "What the web has done for scientific data – and what it hasn't*", Advances in Web-Age Information Management, 1-7.

Carmona, S. (2004). Accounting history research and its diffusion in an international context. *Accounting History*, *9*(3), 7-23.

Chandar, N. and Bricker, R. (2002), "Incentives, Discretion and Asset valuation in Closed-End Mutual Funds", Journal of Accounting Research, Vol. 40 No. 4, pp. 1037–1070.

Chandar, N., & Bricker, R. (2002). Incentives, Discretion and Asset valuation in Closed-End Mutual Funds. *Journal of Accounting Research*, *40*(4), 1037–1070. http://doi.org/10.1111/1475-679X.00081

Charlotte, Barlow (2016). Documents as 'risky' sources of data: a reflection on social and emotional positioning – a research note, *International Journal of Social Research Methodology,* 19:3, 377-384, DOI: 10.1080/13645579.2015.1051797

30

Cheng, Hui G. and Phillips, Michael R. (2014). Secondary analysis of existing data: opportunities and implementation, *Shanghai Archives of Psychiatry*, 2014, Vol. 26, No. 6

Choi, J. and Wang, H. (2009), "Stakeholder relations and the Persistence of Corporate Financial Performance", Strategic Management Journal, Vol. 30 No. 8, pp. 895–907.

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997).Systematic reviews: synthesis of best evidencefor clinical decisions. *Annals of internal medicine*, *126*(5), 376-380

Couture, C. and Ducharme, D. (1998), "Research in Archival Science : A Status Report", Archives, Vol. 4 No. 3–4, pp. 11–38.

Crossan, M.M. and Berdrow, I. (2003), "Organizational Learning And Strategic Renewal", Strategic Management Journal, Vol. 24 No. 11, pp. 1087–1105.

Crossan, M.M., Lane, H.W. and White, R.E. (1999), "An Organizational Learning Framework : From Intuition To Institution", Academy of Management Review, Vol. 24 No. 3, pp. 522–537.

David, P., Yoshikawa, T., Chari, M.D.R. and Rasheed, A.A. (2006), "Strategic investments in Japanese corporations: Do foreign portfolio owners foster underinvestment or appropriate investment?", Strategic Management Journal, Vol. 27 No. 6, pp. 591–600.

Decker, S. (2013), "The silence of the archives: business history, post-colonialism and archival ethnography", Management & Organizational History, Vol. 8 No. 2, pp. 155–173.

Decker, S. (2014), "Solid intentions: An archival ethnography of corporate architecture and organizational remembering", Organization, Vol. 21 No. 4, pp. 514–542.

DeFond, M. and Zhang, J. (2014), "A review of archival auditing research", Journal of Accounting and Economics, Vol. 58 No. 2, pp. 275–326.

Feeney, K. (1999), "Retrieval of Archival Finding Aids Using World-Wide-Web Search Engines", The American Archivist, Vol. 62 No. 2, pp. 206–228.

31

Gioia, D.A., Corley, K.G. and Hamilton, A.L. (2012), "Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology", Organizational Research Methods, Vol. 16 No. 1, pp. 15–31.

Gracy, D. B. (1994). Columbus revisited: the status of archival research around the world in 1992. *Archivum*, *39*, 520–525.

Graffin, S.D., Carpenter, M.A. and Boivie, S. (2011), "What's All That (Strategic) Noise? Anticipatory Impression Management In Ceo Succession", Academy of Management Journal, Vol. 32, pp. 748–770.

Granovetter, M. (1985), "Economic action and social structure: The problem of embeddedness", American Journal of Sociology, pp. 481–510.

Harris, V. (2002), "The archival sliver: power, memory, and archives in South Africa", Archival Science, Vol. 2 No. 1–2, pp. 63–86.

Huber, G.P. and Power, D.J. (1985), "Retrospective Reports of Strategic-Level Managers: Guidelines for Increasing Their Accuracy", Strategic Management Journal, Vol. 6 No. 2, pp. 171–180.

Huber, G. P., & Power, D. J. (1985). Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. *Strategic management journal*, *6*(2), 171-180.

Jacques, R.S. (2006), "History, historiography and organization studies: The challenge and the potential", Management & Organizational History, Vol. 1 No. 1, pp. 31–49.

Jones, G. and Khanna, T. (2006), "Bringing history (back) into international business", Journal of International Business Studies, Vol. 37 No. 4, pp. 453–468.

Ketchen, Jr., D.J., Ireland, R.D. and Baker, L.T. (2012), "The Use of Archival Proxies in Strategic Management Studies : Castles Made of Sand ?", Organizational Research Methods, Vol. 16 No. 1, pp. 32–42.

32

King, G. (2007), "An Introduction to the Dataverse Network as an Infrastructure for Data Sharing", Sociological Methods & Research, Vol. 36 No. 2, pp. 173–199.

Koretz, S. and Lee, G. (1998), "Knowledge management and drug development", Journal of Knowledge Management, Vol. 2 No. 2, available at:https://doi.org/10.1108/13673279810249387.

Latour, B. and Woolgar, S. (2013), Laboratory Life: The Construction of Scientific Facts, Princeton University Press.

Layder, D. (1993). *New strategies in social research: An introduction and guide*. Polity Press.

Lehrer, K., & Paxson, T. (1969).Knowledge: Undefeated justified true belief.*The Journal of Philosophy*, 225-237.

Levinthal, D.A. and March, J.G. (1993), "The Myopia of Learning", Strategic Management Journal, Vol. 14, pp. 95–112.

Lievrouw, L.A., Rogers, E.M., Lowe, C.U. and Nadel, E. (1987), "Triangulation as A Research Strategy for Identifying Invisible Colleges among Biomedical Scientists", Social Networks, Vol. 9, pp. 217–248.

MacCallum, R. (1998), "Commentary on Quantitative Methods in I/O Research", The Industrial-Organizational Psychologist, Vol. 35 No. 4, pp. 1–24.

MacQueen, K.M. and Milstein, B. (1999), "A Systems Approach to Qualitative Data Management and Analysis", Field Methods, Vol. 11 No. 1, pp. 27–39.

Maines, L.A. and Wahlen, J.M. (2006), "The nature of accounting information reliability: Inferences from archival and experimental research", Accounting Horizons, Vol. 20 No. 4, pp. 399–425.

Mckemmish, B.S., Gilliland-swetland, A. and Ketelaar, E. (2005), "' Communities of Memory ':
Pluralising Archival Research and Education Agendas", Archives and Manuscripts, Vol.
33 No. May, pp. 1–15.

Meyer, E.T. and Schroeder, R. (2009), "Untangling the web of e-Research: Towards a sociology
of online knowledge", Journal of Informetrics, Vol. 3 No. 3, pp. 246–260.

Miles, M.B. and Huberman, A.M. (1994), Qualitative Data Analysis: An Expanded Sourcebook,
Sage.

Mintzberg, H. and McHugh, A. (1985), "Strategy Formation in an Adhocracy", Administrative
Science Quarterly, Vol. 30 No. 2, pp. 160–197.

Pattuelli, M.C. and Miller, M. (2015), "Semantic network edges : a human- machine approach to
represent typed relations in social networks", Journal of Knowledge Management, Vol.
19 No. 1, available at:https://doi.org/10.1108/JKM-11-2014-0453.

Payne, S.C., Finch, J.F. and Tremble, T.R. (2003), "Validating Surrogate Measures of
Psychological Constructs : The Application of Construct Equivalence to Archival Data",
Organizational Research Methods, Vol. 6 No. 3, pp. 363–382.

Peters, C.S. (2010), "Searching for 'Spitzer' In the Archives: Drawing New Conclusions from
the 'Troopergate' Scandal Using Archival Records and Social Network Analysis",
Society of American Archivists – 2009 Research Forum.

Pettigrew, A.M. (1985), "Contextualist Research and the Study of Organizational Change
Processes", Doing Research That Is Useful in Theory and Practice, pp. 222–273.

Rowlinson, M., Hassard, J. and Decker, S. (2014), "Research Strategies For Organizational
History : A Dialogue Between Historical Theory And Organization Theory", Academy of
Management Review, Vol. 39 No. 3, pp. 250–274.

34

Ruhleder, K. (2000), "The Virtual Ethnographer : Fieldwork in Distributed Electronic Environments", Field Methods, Vol. 12 No. 1, pp. 3–17.

Storga, M., Mostashari, A. and Stankovic, T. (2013), "Visualisation of the organisation knowledge structure evolution", Journal of Knowledge Management, Vol. 17 No. 5, pp. 724–740.

Welch, C. (2000), "The archaeology of business networks : the use of archival records in case study research", Journal of Strategic Marketing, Vol. 8 No. 2, available at:https://doi.org/10.1080/0965254X.2000.10815560.

Williams, T.A. and Shepherd, D.A. (2015), "Mixed Method Social Network Analysis: Combining Inductive Concept Development, Content Analysis, and Secondary Data for Quantitative Analysis", Organizational Research Methods, pp. 1–31.

Yakel, E. (2003), "Archival representation", Archival Science, Vol. 3 No. 1, pp. 1–25.

Yin, R.K. (2003), Case Study Research: Design and Methods, Applied Social Research Methods Series, Third., Vol. 5, Sage Publications, Inc., available at:https://doi.org/10.1097/FCH.0b013e31822dda9e.

Zhang, W. and Watts, S. (2008), "Online communities as communities of practice : a case study", Journal of Knowledge Management, Vol. 12 No. 4, available at:https://doi.org/10.1108/13673270810884255.

Zwijze-koning, K.H. and Jong, M.D.T.D.E. (2005), "Auditing Information Structures in Organizations : A Review of Data Collection Techniques for Network Analysis", Organizational Research Methods, Vol. 8 No. 4, pp. 429–453.
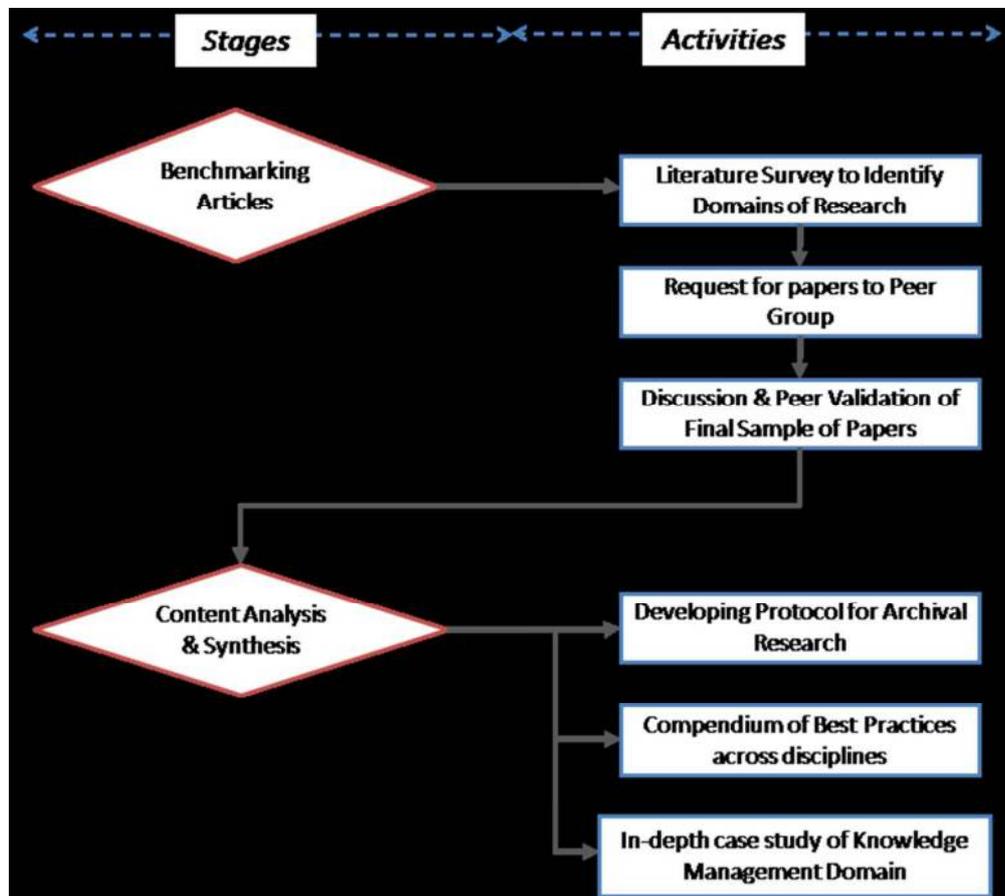
35

Table 1: Illustrative Archival research studies in Knowledge Management

#	Author, Date	Type of Archival data used	Method adopted for analyzing the data	Research question addressed by the study
1	Crossan and Berdrow (2003)	8 years worth of corporate data such as presentations and speeches by top management, annual reports, corporate budgets, sales brochures, union agreements, internal memos and internal restructuring reports. Also External media reports.	These were supplemented with interviews of the top management and analysed using the single case study method of Yin (2003) and Miles & Huberman (1994).	They operationalised and validated their 4I framework (Crossan, Lane, & White, 1999) of organizational learning as a dynamic process, with this case study.
2	Storga, Mostashari, and Stankovic (2013)	5000 email messages exchanged between 2008 and 2012 were retrieved from the internal database of an INGO (International Non-Governmental Organization). These emails pertained to knowledge exchange between members of specific development programs and	Recursive analysis of email interactions was done, supplemented by analysis using social network analytic and visualization softwares to identify knowledge networks and their evolution across time and space.	They demonstrated that (1) the content structure of electronic knowledge networks within the INGO exhibited non-hierarchical and decentralized structure of the individuals contributing to the knowledge discussion boards., (2) that the language of the organization evolves along with the structure of knowledge and (3) the exploration of the knowledge conversion mechanisms from

		geographies.		tacit to explicit., from individual to the group and from informal groups to the whole organization. They also made valuable methodological contributions in: (1) demonstrating coupling of content analysis and social network analysis methods, and (2) Probabilistic modeling possibilities with dynamic data.
3	Aribi and Dupouët (2015)	Extensive archival data, both public (87 documents) and private in nature; collected from 3 firms, similar to those mentioned in Crossan & Berdrow, 2003.	Archival data was triangulated with 23 interviews of top to junior managers and information collated by qualitative analysis with NVivo content analytic software.	The investigated the relationship between firm's absorptive capacity and type of knowledge outcome or innovation output that was expected. They found that firms that targeted "new-to-the-firm" innovations tended to favor the use of social capital; whle those that aimed at "new-to-the-world" innovations tended to rely more on organizational capital. The link with social capital was mediated by speed and reactivity while the link with organizational capital was affected by complexity

			and time-length.
4	Zhang and Watts (2008)	7,853 messages downloaded from the informal online community of software developers, hosted in company database, during a six week period	Qualitative content analysis of the messages	Using the practice-and-identity framework, they showed the existence of communities of practice (CoPs) among online communities and explored how organizations might better utilize these online social structures for their own knowledge management practice
5	Pattuelli and Miller (2015)	50 digitized interview transcripts were collected from special archives spread across the USA. An additional database of all musicians details was created for this paper from external data on the web.	The authors developed semantic software to extract information from the interviews, then referenced these against the database they created and finally used their own visualization software to make the networks.	The aim was to establish the social network connections among musicians in the context of Linked Jazz, a project that applies Linked Open Data technology to oral histories and thus documents jazz history. The authors showed that their applications may be used to effectively link the aforesaid project to the external web and thus streamline the historical documentation program.