




Please cite the Published Version

Cronin, Bruce, Perra, Nicola, Rocha, Luis Enrique Correa , Zhu, Zhen , Pallotti, Francesca, Gorgoni, Sara, Conaldi, Guido  and De Vita, Riccardo (2021) Ethical implications of network data in business and management settings. *Social Networks*, 67. pp. 29-40. ISSN 0378-8733

DOI: <https://doi.org/10.1016/j.socnet.2020.09.001>

Publisher: Elsevier BV

Version: Accepted Version

Downloaded from: <https://e-space.mmu.ac.uk/634308/>

Usage rights:  [Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Additional Information: © 2020. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

Ethical implications of network data in business and management settings

Bruce Cronin¹, Nicola Perra, Luis Correa Da Rocha, Zhen Zhu, Francesca Pallotti, Sara Gorgoni,

Guido Conaldi, Riccardo De Vita

University of Greenwich Centre for Business Network Analysis

1. Corresponding author.

Mail: University of Greenwich, Park Row, London SE10 9LS, United Kingdom.

Email: b.cronin@greenwich.ac.uk

Ethical implications of network data in business and management settings

Abstract

Reflecting on the compilation and analysis of a range of network datasets drawn from our own work and some prominent examples, we consider the ethical challenges in dealing with network data in business and management settings. We argue that the managerial processes that characterize such settings introduce particular ethical sensitivities in the stages of commissioning and research design, and when collecting, analyzing and reporting network data. These sensitivities arise from the imperatives of business, motivations for commissioning network analyses and the legal authority that managers have over employees. We argue that ethical considerations are much more pervasive in business and management network research than in many other fields.

In this contribution, we present a range of ethical challenges in network research in business and management settings that arise at several stages of the research process. For each issue identified, we describe the ethical problem and propose mitigation remedies. From this reflection, we suggest guidelines for other researchers to consider when designing research projects in this application area.

Introduction

As clearly articulated by Borgatti and Molina (2005), social research undertaken in organizations is not simply a matter between researcher and individual participants but necessarily involves a third party, the organization itself, and thus the relationships between the individuals and the organization, and potentially additional stakeholders. In this situation, a researcher's ethical duty to promote beneficial outcomes and to prevent harm to participants may be undermined by the actions of the additional parties towards the participants as a result of the research.

This is especially true in consulting agreements, where those who pay dictate the conditions (Baird, 2014). Such agreements are becoming increasingly common in the neoliberal context, which is driving greater specialization and flexibility in workplaces (Feighery, 2011). In light of these trends it is important to consider ways in which academics might productively engage with those outside of academia in principled and informed ways.

Ethical guidelines are well established in many fields, including social network analysis.

Examples include participants' rights to anonymity, confidentiality, data protection and privacy, voluntary participation and informed consent. Applications of ethical guidelines are commonly encountered in studies of social media, education and health (Moreno et al., 2013; Mannheimer et al., 2016; Wilkinson and Thelwall, 2011; Wade, 2007). Indeed, ethical considerations and guidelines may constrain data analysis, as, for example, in the network study by Kitts et al (2017), where data on individual patients being transferred between hospitals were collected and analyzed. While patient-specific information, such as health conditions or income, was available it could not be used in the analysis because of data protection and privacy restrictions.

The ethical challenges that organizational research pose in business settings, are understudied and are important to examine. We argue that ethical considerations are much more pervasive for organizational network analysis in in this setting than in many other fields and require systematic consideration throughout a research project, well beyond simply seeking approval of an Internal Review Board or Research Ethics Committee. Because of the major societal impact of large businesses and the imperatives driving managers' commissioning and use of research, there are additional socio-ethical implications to be considered. For example, as noted by Hollenbeck and Jamieson (2015), human resource management can benefit from social network analysis when information of how employees interact with each other, especially informally, becomes available. However, the sensitive data collected can also be misused in deliberate misinterpretation of findings or discrimination against certain employees. As another example, the financial industry routinely harvests data from social media to improve credit scoring accuracy. However, data collection without explicit user consent, potential discrimination by scoring algorithms and 'score engineering' by applicants for financial services are prominent ethical issues (Wei et al., 2016).

In most social research, anonymous collection of data minimizes risks of harm to individual participants as the researcher does not know the identity of any individual respondent. Even so, combinations of participant attributes can reveal particular identities where the incidence of some attributes is low. For organizational research, because organizations generally contain small numbers of people and these know each other relatively well, the risks of re-identification from anonymous data are higher.

This is a much greater issue for network research as fully anonymous data collection is typically not possible; other than in ego-centric networks, specific identities are needed in order to capture

data on relationships among individuals. The focus is on the specific relationships between Mary and Jo and between Fred and Jo, rather than some characteristics of most Xs or Ys.

Confidentiality via anonymization of participant identities, replacing names with codes when reporting results, provides some mitigation but the risks of re-identification based on the pattern of relationships remains high (Borgatti and Molina 2003). In addition, informed consent is challenging because social analytic techniques are much less familiar than standard survey research, so participants are unlikely to know the extent to which the information they provide can provide social insights through triangulation of responses. This is even more-so in situations where interconnectivity amplifies the effect of a key network position, such as a critical information bottleneck or a position able to cascade information or misinformation. Not least, in some contexts people are more inclined to disclose someone else's sensitive information than their own (Fishburne, 1980; Rossier, 2010).

Further, Borgatti and Molina (2005) argue that in organizational settings, managers will typically only provide access to the research site if they are provided with the managerially valuable unanonymized data. They thus advocate the use of a management disclosure contract (MDC), specifying what data managers will have access to and the scope of decision-making as a consequence of the research; and a truly informed consent form (TIC), that provides participants with the content of the MDC and examples of data outputs. The latter, particularly outputs in the form of sociograms, is important for informed consent as there is limited awareness of the extent of information that can be obtained from the triangulation of the responses of multiple participants.

In this paper, we argue that the specific context of business presents additional ethical challenges to those commonly identified by scholars in the social sciences and network research, and in particular to the ones discussed by Borgatti and Molina (2003, 2005) with respect to organizational network research. The business context also has implications for researchers' engagement with big datasets drawn from business models aimed at capturing and aggregating massive amounts of personal, relational and behavioral data from customers without their explicit knowledge or consent (Townsend and Wallace, 2016).

We review the ethics of organizational network analysis in the context of business and present some considerations from our own varied experiences applying social network analysis in this area. We first examine the characteristics of businesses that make organizational and network research problematic ethically. We then consider ethical problems related to surveillance in the workplace and business models utilizing digital and online data and the challenges for network analysis these both raise. We draw from our own experiences in organizational network analysis, considering each stage of the research process from research design to data collection, data analysis and reporting. We provide small and big-data examples at each stage.

Business context

Social research in organizations involves relationships beyond those between the researcher and respondents. As Borgatti and Molina (2003, 2005) highlight, the organization itself is a party to the research, a critical gatekeeper and often a commissioner. To this we would add external organizational stakeholders, such as customers or partners, whose information may be held. Due to the characteristics of social network research and the generally small and insular nature of organizations, collecting and reporting network data in organizations entails specific challenges.

However, the challenges of organizational research, and network research in particular, are greater still when those organizations are businesses. In the simple case of employment contracts, employers have the authority to compel employees to provide information for managerial use, including for research purposes, undermining the principle of consent. Some organizations, particularly those in health and social care, have ethical codes and procedures governing research projects. Some professional occupations are bound to codes of practice by their professional accrediting bodies. But most businesses are rarely limited by such codes or procedures.

Even in the absence of binding obligations, the concept of consent in a workplace situation remains problematic. An employment contract cannot be assumed to comprise an agreement between two free and independent actors (Laffer, 2005), nor does a researcher enter the workplace free to contract a management disclosure contract or truly informed consent form on an equal basis. The employer-employee relationship depends on the context, particularly job market conditions, the extent of collective bargaining, symbolic and peer group practices and social norms (Greenwood et al. 2006, Sewell 1998, Van Buren 2001). An external researcher enters this unequal situation with his/her own agenda, not intended solely for the benefit of either management or employee (Guillemin and Gillam, 2004), yet generally dependent on the stronger party, the manager, for any access at all. While in consulting research it is granted that the client dictates what is to be researched, this can also happen in academic research where the researcher is dependent on the manager in order to access data, and for this reason may be under pressure to revise the research questions to align them to the managers' agenda. This is more likely in a business setting where enhancing return on investment is paramount and management will not waste either money or time if there is no benefit to be reaped. In addition, the

dependence of the researcher on others to access data is likely to lead researchers to understate the purposes of the research, or even seek to deceive participants, when asking for consent. Because of power imbalances, this is more likely when researchers are seeking consent from employees within a business.

Additionally, businesses are strongly motivated to collect and analyze information that may affect their performance, particularly information about operational, employee and customer activity. As Borgatti and Molina (2005) note, collection and analysis of data is least restricted in a consulting situation, where both the private consultant and the client are businesses with competitive imperatives to get the most out of the data that can be collected. Consequently, social research in businesses is both less likely to protect participant's ethical rights and has potential for greater impact on the participants' lives than in other contexts. For example, social network research within a business may influence the managerial decision making in promoting or firing someone.

Restrictions on the collection and use of information about and from employees comes predominantly from the legal setting in which a business operates. Examples of this are judicial decisions that allow (US) or restrict (EU) employers from monitoring employee email (Chan 2017). Many countries have prescriptions on the storage and disclosure of personal data. Most extensively, the Europe Union's General Data Protection Regulation demands explicit consent to the collection of personal data in most cases, access to the collected data and the opportunity to correct it. However, consent for the collection of personal data from employees is typically bundled with a range of consents required to commence employment. Furthermore, there are numerous documented incidents of business misconduct, including lying and abusive behavior, especially among firms in developing countries and those operating internationally, which

compromises efforts to follow standards and ethical guidelines (Ethics & Compliance Initiative, 2016).

Surveillance context

The authority given to employers by the employment contract over personal behavior within the workplace includes the power of personal surveillance to a much greater degree than that to which individuals are subject to in civic life. This situation intensifies ethical concerns regarding data collection via observational or surveillance methods in business contexts.

While employees agree to a great degree of management control and surrender of privacy in the workplace, this does not comprise a complete abrogation of human rights; there remain expectations of dignity, respect and private space at work (Greenwood et al., 2006). As with the basic employment contract, however, the scope of human rights varies with the employment context, with prevailing social norms perhaps being the most influential driver in developed economies.

With widespread surrender of metadata in the course of expanding engagement with online business, there has been a recent rapid erosion of social norms with respect to personal privacy in general. Furthermore, the capture of large amounts of personal data by businesses via ambiguous terms of service is a key determinant of competitive advantage (Zuboff, 2019). Similarly, the dependence on digital technologies indirectly forces users to accept terms of service in order to access certain online services. For example, Google was recently fined by the European Union for forcing smartphone manufacturers to pre-install its search engine app and its own web browser as a condition to install other mobile apps in its Android system. The erosion of social norms concerning privacy has been exacerbated by the burgeoning practice of mass surveillance

of online communications by governments for security purposes coupled with disingenuous efforts to deny the power of metadata capture when exposed. Examples include the mass harvesting of email metadata, street cameras, and the Chinese social credit system (Schläeger, 2015). These developments have weakened social norms and sensitivity to provision of private information, which in turn has likely weakened resistance to managerial access to privacy in the workplace. The extent of surveillance and the uses of data collected are not generally disclosed to employees, subject to explicit policy, nor third-party audits, heightening the need for special sensitivity when making use of such methods in business contexts (Ball, 2010; Bloom, 2019).

As with the concept of consent in the workplace discussed above, the ethics of workplace surveillance depends on the power relationship between employer and employee. In situations of low employee power, research facilitating workplace surveillance is less justifiable than in other situations. Even in highly developed countries such as the UK, only a minority of employees feel able to challenge uncomfortable digital surveillance (TUC, 2018). Such issues are particularly critical in developing countries where power relations between management and employees are stronger due to economic and cultural factors that make employees, particularly unregulated workers, more vulnerable. Yet, because researchers are normally dependent on managers for data access, they are likely to be largely insensitive to such power imbalances. For social network researchers in business and management contexts, this adds a further responsibility, as the triangulation techniques inherent in the methods provide much greater surveillance capabilities than traditional social science techniques.

The value from triangulating network data also drives the extension of surveillance to customers as a business model; there is commercial value in knowing customers' social behavior. The imperative to extend the competitive value of this by combining different sources of personal

data, comes from the character of this information as a ‘network good’; its value increases with volume. Competition in network goods markets are characterized by intense races to gain a critical mass of market share, as once a ‘tipping point’ of volume is reached, the vast bulk of the market tends to switch to the dominant firm (Economides, 1996). This is illustrated by the seemingly impenetrable market share of Google, Facebook and Amazon in major segments of personal information. These firms have been repeatedly criticized and fined for harvesting, combining and reselling large amounts of personal data without explicit permission, widely harnessing social network analytic techniques to do so. The scale of the benefits from being first to reach a tipping point exacerbates the normal imperatives of business to cut ethical corners and bend the rules to the point that gross intrusions of personal privacy and opportunism are common. Cambridge Analytica’s attempt to use Facebook data to amplify personal prejudices providing a prominent exemplar (Isaak and Hanna, 2018). Researchers making use of data collected by businesses in this manner, thus need to be very sensitive to the ethical conditions under which it is collected.

Ethical issues in commissioning

Ethical considerations commence with the research question. For example, organizational network studies often aim to understand the impact of structural factors on outcomes of interest such as job satisfaction, task performance and leader reputation (Mehra et al., 2006; Sykes and Venkatesh, 2017). Such research questions require detailed data collection at individual level, such as subjective judgments of peers, which may be sensitive in nature, such as trust (Lusher et al., 2012), negative ties (Marineau et al., 2016), and gossiping (Ellwardt et al., 2012), and allow managerial actions which may be subject to misinterpretation or even abuse.

Academic researchers typically develop research questions by matching gaps in the literature with opportunities to access data. In business and management research, data are accessible only with the cooperation of business gatekeepers, either for access to a site or by providing data to a database. Viable research questions already normally embody a tension between the needs of the research and those of the business. At the stage of research question formulation, it is therefore important to consider the ethics of the question being asked. Should we be addressing this question? Who will potentially benefit, who will potentially be harmed? In business and management situations, the answers to these questions are greatly shaped by the contextual features discussed above. Given the asymmetric power between managers and employees, is it ethically justified to study ways in which employees organize to increase their power within a business? Given the power of social network analysis to add to workplace or social surveillance, will the results of a study transgress workplace or social norms on surveillance acceptability? For academic researchers accessing a business research site via a consultancy assignment, what are the potential conflicts of interest in the researcher's dual identity as researcher and consultant?

In the business network research projects we have chosen to pursue, without explicit discussion of the issue we may be implicitly sensitive to these considerations, as we have tended to pursue inter-organizational more than intra-organizational networks. Where we have taken a more intra-organizational focus, we have pursued projects with clear benefits for participants, such as career progression for minority groups or enhanced collaboration among middle managers. Organizational restructuring projects invited greater scrutiny. In one project, measuring the impact of collaboration from the introduction of a knowledge management system, we canvassed ordinary staff members of the organization in advance and proceeded only once we were assured that the change was broadly welcomed.

Our experience of commissioning of social network research in organizational settings suggests there is an overestimation by clients in the effectiveness of the technique. This may arise from the ‘pitch’ typically used to gain access to a research site. Clients seek clear examples of the technique and clear deliverables from providing access. This leads both researchers and commissioners to converge on exemplar cases, where there is a clearly defined research problem and where social network analysis provides a clear solution to the problem, possibly illustrated by network visualizations. Cross and Parker (2004) provide several such exemplars. In practice, however, the results of social network analysis are normally more ambiguous than the exemplars suggest, and the results lie in the nuances of network metrics with recommendations short of clear-cut solutions. This can be frustrating for commissioners seeking clear, actionable outcomes, particularly in business contexts, and difficult for researchers, pressured to promise business-ready deliverables. In one such case, analyzing collaborative interactions in an industry conference, we were repeatedly asked by a commissioner to collect additional data and repeat our analysis until we were able to convince him that his assumption about the central role of some high-status individuals was not supported. Such ‘scope creep’, is frustrating to both parties, adds costs to the researcher and undermines ongoing relationships.

In our experience of business consulting situations, the champions of social network research within organizations are typically ambitious, early adopters of new techniques, who open the door to social network research but can also close it prematurely. We have several experiences of multi-wave research plans where the first wave of data collection was completed but when it came to presentation of initial findings and preparations for the second wave, the champion had moved on to another organization, leaving no one behind sufficiently motivated to understand

the initial purpose or support continuation of the research. Again, this is particularly difficult in business contexts, where a clear return on investment is expected.

The dual identity as a researcher and a consultant also creates a dilemma that the social network research may be of high intellectual value to an academic career but of little or no value to the client or vice versa (Morton, 1999). An example of the latter is what Faria (2001) calls the ‘consultancy disease’, which is particularly common in developing countries where academic rewards are less attractive financially than private consultancies, but where there is a clear incentive to enter academic life as a way to access external earnings because of the reputation attached to a university position, the main channel to obtain consultancy work.

Borgatti and Molina’s (2005) interrelated management disclosure contract (MDC) - truly informed consent form (TIC) is valuable in focusing realistic and workable outcomes, though we have found managers reluctant to fully disclose research aims in the TIC. The management disclosure contract also needs to explicitly consider the full range of stakeholders who will have access to or provide data. These may include regulators, clients and customers of the firm. Where management provides data collected from clients or customers for analysis, it is unlikely that the customers, when providing the data, were aware of the extent of information about them that can be derived from triangulation of relational positions and agglomeration of data sources. Where we have made use of such big datasets, we have been very careful to restrict reporting to very abstract levels.

The key recommendation for remedying commissioning problems is greater explicit specification of procedures and deliverables, particularly greater disclosure of the limitations of social network analysis, albeit at the risk of reduced access to research sites. Informed consent is

enhanced if all participants, including organizational gatekeepers are aware that the technique is not a ‘killer app’, yet can still provide important insights into phenomena not normally on the radar of business. The informed consent of all stakeholders needs to be considered, not just the immediate gatekeepers and individual participants. And the conditions under which big datasets are compiled need to be considered when reporting is planned.

Ethical issues in research design

The ethical challenges of social network research, heightened in business and management settings, demands careful attention to these in the research design stage. Research design issues of ethical importance in this context include the framing of the relationship between the researcher and the research site and the implications of the sensitivity of many network methods to variations in data. Methodological choices such as sampling frequency affect the identification of network contacts, potentially generating false positives or false negatives and thus resulting on unreliable network structures. Practical issues, such as available resources, population size, or rights, limit sampling and consequently the reliability of collected data.

Because of the sensitive and personal nature of data collected for social network analysis, the framing of the relationship between the researcher and the research site has a great impact on the data collection process. The researcher needs to stand sufficiently distant from stakeholders to diminish mistrust but sufficiently close to encourage disclosure. Where traditional positivist research advocates a dispassionate and neutral stance by a researcher to facilitate data collection from a research ‘subjects’, this creates a subordinate position for the subject as an instrumental object of the researcher (Oliver, 2010). This objectification is readily expressed in practices ranging from insensitivity to arrogance or much worse and invites the subject to resist

objectification through satisficing strategies, such as offering information the subject believes will satisfy the interrogator, while minimizing disclosure of information that may cause the subject harm. Research in business contexts, where employees are subjects of command, classically generates such dynamics (Covaleski et al., 1998; Alvesson, 2003). Additionally, with its quantitative orientation and roots in positivist sociology, social network analysis is vulnerable to this framing (Abbott, 1988; Emirbayer and Goodwin, 1994), as reflected in frequent discussion of resistance to personal disclosure in questionnaire design

More fruitful data collection comes when the relationship between data collector and provider is framed more equitably, as researcher-respondent, interviewer-interviewee, research lead - research participant (Graham et al., 2007; Oliver, 2010). This framing continually asserts the interests of data providers when planning, collecting and analyzing data and provides some reassurance to them about the use of their data, supporting fuller disclosure. The framing of data providers as ‘participants’, rather than ‘subjects’, implies involvement in research design, common in qualitative and ethnographic research where personal disclosure risk is greatest. This may also be more appropriate for personal relationship data for social network analysis.

Involving participants in research design can provide wider insight into the research setting to refine research questions and improve data collection and alert researchers to sensitivities of participants (Oliver, 2010). However, this framing does not entirely resolve the relationship dilemma as many issues of power in decision-making remain and research leads may struggle to fully engage participants in the project because of differing priorities. Elsewhere in this special issue, D’Angelo and Ryan (2019) discuss the performative aspects of these interactions and the value of reflexive approaches to social network research.

We have tended to frame our own questionnaire-based social network research with the terms ‘researcher-participant’. This has provided sufficient distance and closeness to avoid explicit objection to questions asked but is deployed more in the spirit of ‘interviewer-interviewee’ than any involvement in the research design, which participants appear to view outside their competence (see similar observations in Graham et al., 2007).

Social network analysis occupies an atypical position in the polarity of quantitative and qualitative research. Quantitative research typically involves systematic examination of large amounts of detailed data within tightly specified categories while qualitative research typically derives categories from nuanced wide-ranging examination of phenomena. But social network analysis is highly reductive, reducing complex, often intangible, multifaceted qualitative interpersonal relationships to a set of 1s and 0s on a single dimension; it then undertakes very detailed quantitative analysis of these data. These circumstances risk the presentation of spurious precision in substitution for the typically more nuanced and ambiguous qualitative data underpinning the analysis. While this is a risk in all quantitative social research, it is more problematic in social network analysis as the conventions in identifying, collecting, aggregating and categorizing the underpinning data are much less developed than in many other fields (Pallotti et al., 2020), although progress has recently been made in qualitative methods for network analysis (D’Angelo and Ryan 2019). Issues for network research include network boundary specification (Fitzgerald et al., 2002; Laumann et al., 1983; Wasserman and Faust, 1994), data aggregation (Quintane et al., 2014), temporal resolution (Rocha et al., 2017), network completeness and subjective bias (Fischer, 1982).

The critical issue here is that methodological choices affect the identification of certain network contacts which in turn may over- or under-estimate network connectivity. The frequency and

length of data collection may generate (false positives) or miss (false negatives, censorship) interdependences that result on unreliable network structures. While false positives may indicate inexistent social or professional interactions (e.g. email communication between fraudsters), false negatives may hide information, breaking chains of interactions at both the employees (e.g. power relations via mid-managers) or businesses (e.g. missing intermediates in supply chains) levels. Researchers must be aware on the limitations of data collection for dynamic networks since the structure may be different at different times. We have found structural patterns from temporal network data analysis of human communication and physical interactions highly dependent on the chosen temporal resolution (Conaldi et al., 2012, Rocha et al., 2017).

The specification of network boundaries has ethical implications. In research on interlocking directorships among firms, when considering the social impact of concentrations of corporate power, we have struggled to identify the appropriate scale at which to undertake analysis. Overly narrow network boundaries overemphasize the prominence of some firms; overly wide boundaries dilute the importance of such firms (Cronin, 2017). The extent to which the boundary specification captures the object of study determines the extent to which structural features of the network accurately represent the actual social situation. For example, incomplete sampling makes questionnaire-based social network research susceptible to respondent bias in question interpretation; in work situations there is considerable variation in interpretation of relationships such as ‘friendship’, ‘advice’ and ‘trust’ (Cross and Parker, 2004). On the other hand, snowball sampling approaches routinely overemphasize network connectedness. So, where this sampling cannot be avoided, as in very sparse organizational settings like rural business communities or hard-to-reach populations (e.g. unregulated workers, illegal immigrants or elite investors), we have restricted our analysis to ego-centric networks. (Rocha et al., 2010).

The sensitivity of network analysis to such design has ethical implications as spurious or erroneous identification may generate an unwarranted interpretation that may impact on participants' lives. This is particularly so in social network research where small design errors may wrongly place particular individuals in critical network positions. The consequences of this in business management consulting, are even more serious than in the academic setting, because the purpose of the network research is to aid decision-making that will have as a result an effect on the lives of employees. Cross and Parker (2004) provide an anecdote of a manager interpreting peripheral positions in an organizational network as redundant roles, overlooking other dimensions of their work as externally-focused sales agents.

Such problems can be mitigated by sensitivity to the researcher-participant relationship and more rigorous research design, including greater use and reporting of robustness checks on boundaries and missing data than is typical in contemporary organizational network analysis. This includes the reporting of completeness statistics the use of established scales and construct validity checks when constructing survey instruments.

Ethical issues in data collection

The ethical issues in protecting participants in business settings when collecting network data via survey instruments have been broadly discussed in the introductory section above. Well-informed participants, confident of confidentiality, strengthen research through greater engagement and data disclosure (Graham et al. 2007; Singer, 1978). But because of the distinctive focus in social network analysis on individual relationships as the unit of analysis, the scope for impact on individuals as a consequence of collecting network data is heightened in

business contexts where participants typically have a power disadvantage or individuals are unaware that data about them is being collated.

In particular, because false positives or false negatives in network data are amplified in network analysis, the position of an individual in a network can be more readily misattributed than with other research methods. As misattribution of network position can have negative impacts on that individual or others related to them, there is a heightened *ethical*, and not just methodological, duty for researchers to give careful consideration to the ‘hygiene principles’ (Monaghan et al., 2017) in data collection for SNA. These principles are reviewed in detail by Marsden (1990, 2005), who notes that the handling of these are seldom explicitly reported.

As Marsden (1990, 2005) catalogues, there are particular issues for the collection of network data in terms of construct validity, informant accuracy and competence, identity ambiguity, and researcher effects. Surveys, questionnaires, name generators and name interpreters need careful design so that information sought on particular relationships is actually captured. Informants tend to recall frequent, intense and recent relationships over others. Reciprocated reports are closer to observed relationships than unreciprocated reports. Centrally located informants are more competent in reporting relationships over others. Archival and observational data is prone to ambiguous description of subjects of study. Interviewer variation affects name generators more than other forms of surveys, while administration of a network questionnaire after an interview prompts more contextualised responses. Kilduff and Oh (2006) emphasise the importance of contextual information alongside raw network data in interpreting relationships.

Amidst these hygiene principles for network data collection in general, some are particularly salient in business and management contexts. Questionnaire design can be methodologically

problematic in business settings, access to competent informants on inter-firm relationships can be challenging, and identity ambiguity is prevalent in commercial datasets.

Data collection via questionnaires in business settings can be limited by the task structure among participants. This greatly conditions otherwise seemingly independent social relationships; participants have systematically greater interaction with some members of an organizational unit than others. Related to this, even in small-scale roster-based survey research, names presented may differ from those commonly used by members in an organization; we have found that last names may not be well known, nicknames or shortened names may differ from those listed by the HR department.

The quality of inter-organisational network data varies greatly (Zaheer and Usai, 2004). Many inter-organisational studies rely on heroic assumptions on what occurs in a director interlock or co-patenting registration. Where inter-organisational data are derived from questionnaires or interviews, the quality depends greatly on access to informants sufficiently competent to report the salient elements of the relationship under study; more typically informants are convenience sampled. In our experience, the quality of inter-organisational data is enhanced by surveying a variety of managerial roles within each firm.

Identity ambiguity is pervasive in commercial datasets. When using large intra- and inter-corporate network datasets, we have encountered considerable challenges in systematically identifying both corporate entities and those holding managerial roles within these. It is difficult to correctly classify “BP US”, “British Petroleum (Inc)” and “BP Incorporated”, or “John Smith”, “J.S. Smith”, “J. Smith” and “John S. Smith” when they occur in the same dataset. This is an even greater problem in the analysis of co-authorship or citation in scientific networks, where

corroborating information such as middle names or birthdates are rare. The issue is not homogeneously distributed across nationalities and countries; for example, the 100 most common surnames account for about 85 percent of the total population in China (Liu et al., 2012). There is also considerable ambiguity in potentially supporting meta-data such as institutional affiliation; people move between institutions and refer to the same institution in different ways.

Misattribution of connections in these contexts can greatly exaggerate or diminish personal and institutional reputations.

Commercially-produced databases normally make attempts to provide unique identifying schemes, such as ORCID numbers in scientific publication, but we have found such schemes with corporate data to be far from free of inconsistencies (Heemskerk et al. 2018). For small datasets, multiple independent coders and measuring inter-coder reliability are helpful. For larger data sets we make use of standard name matching algorithms, though few of these are optimized for corporate data settings. There are a vast number of advanced algorithms trying to identify duplicates (Deville et al., 2014; Milojević, 2013, Radicchi et al., 2009; Zhang et al., 2013), including crowd-sourcing approaches (Kaur et al., 2012) but this is a dynamic process with problems difficult to avoid (Kagan et al., 2018). We therefore suggest that the best approach here is to favor reliable data sources, for example, data that goes through internal self-regulation or filtering (Rocha et al., 2010); make efforts to identify outliers, for example, super-active users; and remove poorly active or dormant users; always recognizing the potential limitations of data due to duplication of users and links.

To reiterate, it is important for network researchers in business and management settings to pay attention to hygiene factors in network data collection because of the potential for errors to have real impacts on participants' working lives. Among these hygiene factors, our experience in

business settings suggests particular attention is needed to the construction of rosters, identification of informed respondents and identity disambiguation.

In addition to the ethical duty to give proper attention to hygiene principles, we have encountered some other ethical challenges in data collection in business and management settings. These concern the way collection of data about third party employees adds to their surveillance, concerns in the use of some observational techniques, the use of data provided for another purpose, and questions of data ownership.

Gathering data from egos about third parties without the third party's consent is troublesome when this discloses personally sensitive information, such as unregulated or illegal work or protected characteristics such as sexuality, political or religious beliefs, or disabilities. But in business situations the asymmetric power of managers versus employees and suppliers versus customers needs attention; collection of data about third party employees or customers adds to their surveillance. In our experience, first special consideration of the social value of collecting this information is needed, erring on the side of not collecting the data. Then, if the data collection can be well-justified, special care and checking is needed to avoid the reporting of alter attributes that potentially allow identification of these third parties through triangulation. In our research on peer interactions in high-risk situations, we have averaged the third parties' traits or opinions (asking this information indirectly through the egos) instead of linking characteristics to specific people. The local structure is lost but privacy is respected (Chen et al., 2018).

In response to the difficulties and costs of collecting network data through questionnaire or survey methods, business managers are usually very open to the use of observational techniques. These include tracking RFID tags, sensing badges, location-tracking apps, engagement with IT

systems logged, or social interactions recorded by video or CCTV (Cattuto et al., 2010; Kim et al., 2012). Despite the methodological attractiveness of accessing complete datasets in this manner, however, researchers again need to consider the power asymmetries in a business and how the data and results are likely to be used. Consent issues abound; it is unlikely that participants will be fully aware of how much individualized data can be collected from such techniques.

Further ethical challenges arise where observation of user behavior in online social networks such as virtual teams is sometimes only possible by adopting false identities, such as bots, in minimal-risk interventions (Boshmaf et al., 2011). Ethically, this is more personally intrusive than collecting self-reports of alters, as the scope for participants' strategic agency is reduced. This is likely why we have found it more difficult to engage respondents with observational techniques than with questionnaire approaches. In our research observing team dynamics, we have found that confidentiality concerns can be mitigated by anonymizing identities before further analysis, such as substituting confidential id numbers, purposefully blurring the face or using software to generate line drawings from images. More generally, ethical responsibilities involve consideration of the surveillance context in the organization with and provision of detailed explanation of the extent of movement data captured and the implications of triangulating this with others' movements at the consent stage. The benefits of deceptive observation in online social networks needs to be very carefully balanced against not just risks to participants but also the impact on the medium such as exposure of weaknesses and other sensitive information, resource usage, distortion of activity statistics, and impacts on user behavior and on external stakeholders such as advertisers (Elovici, 2014).

An underexplored area in research ethics is when data collected for one purpose are used for another purpose. When the original data are collected with informed consent of the providers, it is unlikely that there would be consent to an unimagined and different subsequent use. As discussed earlier, there are strong business imperatives to mine customer behavioral data and triangulate this with other personal data sources. Public funders of research also encourage the exploitation of secondary datasets, with the onus for ethical use placed on the original data-owner. But, as highlighted by the Cambridge Analytica case, checks by data owners of data use are far from systematic, and business imperatives encourage minimization of compliance with regulations. Researchers making use of data collected by businesses in this manner, thus need to be very sensitive to the ethical conditions under which it is collected.

A particularly troubling example we encountered was a commission to understand how advice and friendship networks affected the career progression of ethnic minorities within a large organization. Midway through the data collection, a commissioner mentioned that one subunit in our sample had a reputation for racism, ‘they are very white’, and that they would be particularly interested in what we found here. Racism was not an explicit construct in our design and we were not collecting any data on this *per se*. But it was clear that our network analysis would be interpreted in these terms. We were unsure how to deal with this other than to warn against over-interpreting the findings. In the event, the analysis suggested the networks in this unit supported the careers of ethnic minorities well. The commissioner concluded that the members of the unit ‘knew how to answer the questions’.

A final consideration when collecting data in a business context is the question of data ownership. While data may be collected via an individual relationship between researcher and participant or drawn from an organizational database, the organization always stands as a related

party. Unless otherwise specified, the organization retains unambiguous legal ownership of any data held within its boundaries. Iphofen (2009, p. 38) cites a case where a corporate commissioner claimed ownership of un-anonymized original questionnaires collected from staff. The researchers were torn between a desire for continued access to the research site and the assurances of confidentiality they had given to the staff. They opted to follow what had been recorded in the ethics approval form concerning obligations to the commissioner, at the risk of being sued. In such a case the researcher is unable to meet their commitments to research participants.

In summary, ethical problems in the collection of network data in general are well-known and ameliorating practices well-established. But in business settings, the use of rosters, locating competent informants and identity ambiguity are particularly problematic. In addition, data collection on third parties and observational techniques in general need consideration of the surveillance context of the business and careful consideration of the potential harm relative to benefits. We recommend more attention be given and more systematic reporting provided in the collection of sensitive information about third parties and the use of existing datasets for new purposes. These issues and implications of data ownership need to be well specified at the commissioning stage.

Ethical issues in data analysis

Business imperatives impact on data analysis, raising ethical challenges. A key concern in business-commissioned research is to return managerial-valuable results actionable in a short time period. In this situation there is pressure on researchers to favor pragmatic and fast methods over more thorough ones, and incremental outcomes over more critical investigation. We

consider ethical challenges that arise when combining data from different sources, interpolating missing data, when making choices of analytical technique, and when making culturally-based assumptions.

An analytical area of ethical concern in network research is the combination of datasets, where this amplifies the power of triangulation to provide often new information about individuals, unbeknown to them, and increasing risks of reidentification. Business imperatives motivate telecom companies to examine their subscriber behavior and commission researchers to combine datasets to draw insights. We have used mobile phone Call Detail Records (CDRs) to understand the organization and structure of social communication at scale and over time (Karsai et al., 2014). CDRs record ostensibly anonymous attributes of phone calls, including start time of call, duration, completion status, source phone number and destination phone number. But the combination of these network data with metadata held by phone companies on birth, zip-code of residence and gender allows large parts of the population to be uniquely identified (Sweeny, 2000). Here again, separated consents may regulate and allow the use of each data but may not address the potential and consequences of combined use of the data, that is in fact difficult to predict in advance. Innovative methodologies may even reveal novel relationships or patterns, that are beyond the original consent of individuals. Since the prediction exercise is difficult, in such cases, the ethical burden should be transmitted to researchers and companies, as for example, by reporting averaged instead of individual data. Even without metadata, the risks of re-identification from CDRs is high. The subset of source and destination phone numbers and call duration is enough to build the communication network between users. CDRs contain also information about the antenna(s) supporting both ends of the communication. Leveraging such information, the location of the user initiating or receiving a call/text can be determined. Even

more, by observing long time periods, we can reconstruct the mobility of each individual. The resolution and precision is not that of GPS; the mobility is reconstructed at the level of geographical cells centered on antennas. The size of the cell is function of the number of antennas, typically smaller within densely populated areas (i.e. cities) and larger in rural places. Nevertheless, it is possible to infer information about home, work, other favorite locations and commuting patterns with good precision (Zang and Bolot, 2011). Indeed, Golle and Partridge (2009) showed how a sizeable fraction of working population in the USA can be uniquely identified by knowing home and work location even at small geographical resolution. One might think that the risks of re-identification from complete CDRs can be resolved by providing anonymized mobility traces; in these settings, there is no explicit information about which user originated which trace. However, the nature of human mobility makes this solution unviable; we explore space slowly, we are highly predictable, and mostly unique (Toole et al., 2015). As a consequence, by observing only four different spatio-temporal location points from CDRs we can uniquely identify 95 percent of the individuals (de Montjoye et al., 2013).

Is it possible to reduce risks while advancing scientific knowledge? In practice, anonymity breaches are mostly a concern when the data is published or publicly shared. Many of the papers mentioned above are the result of closed collaborations between few researchers and data providers which are often not disclosed. In such scenarios research is often conducted following the *remote access* or *question-and-answer* models (de Montjoye et al., 2018). The data provider does not want to or legally cannot lose control of the data. Thus, the data are anonymized but never shared or moved from the secure servers of the data provider. Furthermore, outputs are highly restricted and coded in detailed legal agreements. The question-and-answer model adds yet another level of control and protection over the data carefully selecting the type of questions

that can be investigated. While this model drastically limits open-ended and explorative analysis it still allows the extraction of pieces of information which can be used to inform research.

A second, related, ethical issue in analysis is interpolation of missing nodes or edges based on the egonets of known individuals. This is the basis of friend recommendations in social networking platforms such as Facebook and LinkedIn. In short-term business consultancy assignments, we have experienced pressure to interpolate data in this way rather than returning to the field to collect more data. In pursuit of business opportunities, egonet interpolation techniques were used by Cambridge Analytica to infer individual attributes - psychological characteristics and likely voting intentions of individuals – that were used to try to influence individual voting behavior. In a similar manner, there is scope for businesses to pursue price discrimination in online purchasing based on related users' historical behavior and profile data. With the phone data discussed above, Zang and Bolot (2011) found re-identification of anonymous mobility traces could be greatly increased by having access, even partial, to information about the size of the ego-network of each individual. The social dimension introduces another discerning factor that can be leveraged in the re-identification process. Studying these data sets require extra caution because prediction always contain a level of uncertainty and thus risks generating false positives or negatives. Here prediction is based on models and not observations but the ethical consequences are similar to those emerging from mistaken data collection because identifying inexistent contacts or missing existing contacts may lead to misleading interpretations.

A third analytical issue with ethical implications in network analysis is the choice of analytical technique. While the appropriate analytical technique is normally driven by theoretical considerations underpinning the general research design, in many applications of social network

research in the field of business and management network theorization is underdeveloped (Borgatti and Halgin, 2011; Ahuja et al., 2012; Shipilov and Gawer, 2020; Kurt and Kurt, 2020). The ethical risk here is that inappropriate analytical techniques may over- or under-state critical network positions of individuals. The vast majority of studies of the impact of director interlocks on firm performance in the finance literature, for example, employ degree centrality and no other network analytic, whereas, as found in our research, eigenvector and betweenness centrality are more important determinants (Cronin and Popov, 2005). Yet the limited findings in the finance literature have supported regulatory changes to limit the types of board memberships directors may hold. There is a similar issue with clustering or community-detection methods, where choice of technique can lead to radically different groupings of nodes (Fortunato and Hric, 2016). We recommend robustness checks with respect to standard metrics when analysis concentrates on a single network metric.

Finally, sensitivity is needed to the cultural biases of extant business and management network analysis. The central propositions and theories of business and management applications of social network analysis have been developed in US high-tech large business settings, often focused on the social relationships of middle or senior managers, generally white and male. Research on business networks in knowledge management and corporate governance universally applies the same theories directly in non-western settings, sometimes identifying a skill or knowledge ‘deficit’, where the difference may be attributable to a difference in socio-economic or cultural conditions. A promising alternative is the growing study of the distinctive impact of inter-personal ‘guanxi’ relationships in Chinese management (Badi et al., 2017; Burt et al. 2018; Horak, 2018; Lin, 2011). There is also potential to develop and extend an earlier body of work on culturally-distinctive inter-organizational relationships, such as the Keiretsu and Chaebols

found in East Asia (Hamilton, 1996) and nationally distinctive forms of corporate director interlocking (Windolf, 2002; Kogurt, 2012).

Ethical issues arising in the analysis of network data thus require careful consideration. By amplifying opportunities for triangulation of attributes, the combination of datasets and interpolation techniques raise increased risks of inadvertent identification of individual participants, the generation of additional information about them, probably well beyond their expectations when providing the data. This potential needs to be thoroughly assessed in the design stage and incorporated into the consent process when collecting data. When using existing data, spatial and temporal restrictions on the scope of data combination may be needed, and a conservative presumptive approach adopted – should we be doing this at all? Researchers should be wary of relying on a single or very limited set of metrics in network analysis as this may provide a distorted view of an individual’s role in what is normally a multifaceted social situation; robustness in analysis is important ethically as well as scientifically. Lastly, researchers need to consider the generalizability of their analytical constructs, aware of the predominantly American management context in which organizational network analysis has developed.

Ethical issues in reporting and interpretation

In this last stage, we discuss ethical challenges in reporting findings in a business and management context, including problems of re-identification, unskilled interpretation, unauthorized data sharing, and non-disclosure of information important to participants.

The visualization of networks in sociograms is a powerful communication technique, providing the means to present data richly (Tamassia, 2014). A sociogram simultaneously presents micro, meso and macro context, that is, individual node positions and dyadic relationships, regions of

greater or lesser connectivity, and the structure of a network as a whole, and via comparisons, its evolution in time. But for small datasets, as discussed in our research, even when individual names are not reported, it can be quite straightforward to re-identify individuals from the pattern of interaction that is reported, (Linhares et al., 2017). In a small organization, individuals with high degree centrality are likely to be well known. This is even more the case when attributes such as gender, role, or geographic location are incorporated in the visualization. In principle, the issue is less of a problem for big datasets, though care is needed in the visualization of rare network positions or rare combinations of attributes or the use of filtering algorithms. Even at scale, companies may be wary of revealing strategic partnerships, the structure of supply chains, or logistic networks that may become clearly visible to competitors using sociograms.

The problem is greatest in business consulting, where the commissioning is pitched on the communicative power of such visualizations and these are expected delivery outcomes. If the consulting is provided by an academic researcher, there may also be a conflict between research quality (e.g. statistical robustness) and consultancy quality (e.g. ease of interpretation) when the results are presented (Morton, 1999). For small datasets, we recommend the reporting of metrics and use of abstract or exemplar visualizations rather than visualization of the actual data.

Network visualization is highly dependent on the algorithms used to draw the sociograms. The location of nodes is typically arranged to reduce cluttering of links or to emphasize certain structures, for example, communities, clusters, hubs or the spatial dimension (Tamassia, 2014). Consequently, the algorithms used should be reported unambiguously and visualization should be combined with robust statistical analysis to illustrate or indicate patterns. Sociograms are particularly problematic in the context of unskilled interpretation, where the visualization may simply reinforce existing assumptions or prejudices. In consulting situations, we have

encountered extraverted participants ‘seeing’ themselves at the center of exemplar visualizations. Cross and Parker (2004) cite a manager who interpreted peripheral nodes in a network visualization as candidates for retrenchment. Even the reporting of network metrics needs to be well-interpreted; a position of centrality in a network is seldom unambiguously positive, as is often assumed in organizational contexts.

On the other hand, beyond the particular objectives of the research or those commissioning it, researchers have an obligation to communicate the findings to stakeholders in an accessible manner (Oliver, 2010). In business and management situations, where research is often commissioned by an organization’s management, managers and researchers may have different views on what should be communicated, with managers having great power over this. Elsewhere in this special issue Tubaro (2019) reflects on the competing ethical obligations in such situations.

Reporting is more problematic when disclosure of results may have important benefits to participants. While there is ethical value in a conservative presumption in reporting data, there may be benefits in disclosing findings to individuals concerning health or safety, performance analytics or even network analytics. Ethical concerns in social research concentrate predominantly on the avoidance of harm, non-maleficence. But they also include obligations of beneficence, for research to provide a social benefit. These are not restricted to contributions to knowledge in general; they also include specific benefits to the research participants themselves. Thus, there is a duty of reciprocation to those that have provided their data and others in similar situations, particularly to vulnerable populations (Beauchamp and Childress, 2001; Fontes, 1998; Tubaro, 2019). In business and management applications, this suggests particular obligations to employees over management, and in organizational network analysis, sensitivity to the ways

knowledge of social networks in a workplace can enhance power of particular stakeholders. Again, these are issues that need early consideration, specification and agreement at the commissioning stage.

A more general but fundamental ethical aspect related to reporting and interpretation, is one discussed by Baird (2014) with respect to private commissioning of research, where contract law obliges researchers to produce what the employer requires, and therefore where the latter has every right to alter the content of the report despite the researchers' opinion. This is something that needs to be addressed at the commissioning stage, and in particular when negotiating the contract, but which is even more difficult in social network research because of the more pronounced dependency of the researcher on the client, as discussed in the initial section of this paper. In the case of academic consultants this adds an additional layer of ethical challenge. Baird (2014) denounces the fact that it is becoming increasingly common for academics to sell their intellectual property without challenging the substance of the consultancy contracts, thus leading to violations of our fundamental ethics as academics, where scientific freedom is the basis for both the legitimacy and credibility of research.

We have found business consulting assignments a useful means to access very sophisticated business and management situations, to apply recent developments in network analysis and to derive interesting results. But we have been repeatedly frustrated by the channeling of reporting to very narrow business outcomes and contractual restrictions on publication, largely because of insufficient specification of publication rights against 'standard commercial' contracts for fear of losing the access opportunity. This has started to change as we have accumulated experience in this field, our sense of the value we provide has increased, and realizing the time-horizon of business is quite limited. As the results of research assignments often have competitive value for

only six months or so, managers are generally agreeable to publication after a short embargo. Clearly, potential conflicts of interest need to be transparent, so funded consultancy research should be reported as such.

A final, societal ethical challenge is consideration of the position of business in society. While there is a normal presumption of anonymity and confidentiality by default, Iphofen (2009) observes the common practice in business and public policy research to reveal the identity of organizations and key informants. Subject to informed consent, this identification is justified because they are already in the public view, often interviewed and discussed in public media, and are assumed to be less vulnerable than non-elite subjects. Walford (2005) argues that anonymizing research sites elides typically site-specific effects or specific systems of institutional relationships (Nespor, 2000). Tilley and Woodthorpe (2011) see organizational anonymization stifling opportunities to highlight good practice. Others argue that organizations, even if legally constituted as ‘persons’ as corporations are, do not have the same ethical rights as people, particularly where they have considerable power in society (Aldred, 2008; Langlois, 2011) and to do so amounts to a privatization of social knowledge (Nespor, 2000).

We have repeatedly encountered situations where a Research Ethics Committee (REC) has asked for the reporting of ‘Plastics Co.’ or the like rather than the explicit identification of a major multinational against our desire to situate the study more concretely. We have generally resolved this dilemma by asking organisational managers whether they were agreeable to identification of their firm. Often they have been and the REC has agreed, leaving us to confront the issue again with journal editors. But should we be strictly bound to the consent principle if a socially powerful organization objects to be identified? And what about less powerful organisations seeking publicity for their activity? In both cases, how do we properly account for the impact this

may have on organization members? Ethics committees are risk-averse; elsewhere in this special issue, Molina and Borgatti (2019) discuss the dampening effects of the conservative presumptive stance of these “moral bureaucracies” on social science research and social network analysis in particular.

At the reporting stage, then, considerable care is needed in the ways in which findings are reported, particularly in commissioned research. To mitigate risks of re-identification of individuals, network visualizations should be used sparingly, if at all in small dataset situations. Findings need to be well-contextualized in organizational and business environments to avoid unskilled interpretation and unwarranted managerial action. And issues of data protection and warranted disclosure need to be well-anticipated at the commissioning stage and when seeking institutional approvals.

Conclusion

Researching social networks in business and management settings entails particular ethical considerations beyond those of organizational research and social network analysis in general. The business and managerial processes that characterize these settings introduce particular sensitivities arising from the legal authority that managers have over employees and the imperatives of business.

Yet, the vast majority of research on the ethical implications of social network research are in the areas of education, social media and health. In these disciplines, progress has been made towards agreed principles of best practices, which instead are lacking in business and management.

This paper aims at filling a gap in the literature by suggesting a systematic approach to identify and address the ethical challenges that emerge along the chain of activities involved in undertaking research specifically in the area of business and management. Our major recommendations that follow, are presented in checklist form in Table 1.

In the commissioning stage, we find then need for more explicit specification of procedures and deliverables, including the disclosure of the limitations of social network analysis to manage what are often high expectations. Informed consent of all stakeholders also needs greater consideration, particularly the conditions under which large datasets are compiled from user-provided data.

In the research design stage, careful consideration and reporting of network and boundary definition is needed to avoid errors that can wrongly generate unwarranted interpretation and impact on individuals' lives. We recommend greater use and reporting of robustness checks, reporting of completeness statistics, the use of established scales and careful attention to construct validity when constructing survey instruments.

In the data collection stage, while there are established practices to support rigorous collection of network data via survey instruments, we find careful attention is needed in the design of questionnaires, observational techniques need consideration of the surveillance context of the business. Issues of identity ambiguity can be mitigated by the reporting of inter-coder reliability statistics, standardized name-matching algorithms and error rates. We recommend more attention be given and reported on the collection and protection of sensitive information about third parties and the use of existing datasets for new purposes. Data integrity may be particularly problematic

in poor or unregulated settings. These issues and the implications of data ownership need specification in commissioning.

In the data analysis stage, we find ethical issues arising when combining datasets and interpolating missing data and selecting analytic techniques. These issues can be mitigated by careful consideration of the risks of triangulation when combining datasets, spatial and temporal restrictions when mining existing datasets and robustness checks on network metrics. We would welcome greater critical consideration of the, often culturally-biased, analytical constructs used in much organizational network analysis.

Finally, we find reporting findings raise ethical challenges in business settings, where strong incentives exist to leverage extensive data for business performance purposes, with potential important impact on the research participants. To mitigate risks to participants we find particular care is needed in the deployment of network visualizations, findings need to be well-contextualized when presented and issues of data protection and warranted disclosure need to be well-anticipated.

In addition to the specific issues discussed above, some general issues call for further reflection and research. First, and related to the design and data collection stage, is the importance of the cultural context, and the fact that ethics is not perceived in the same way in different business settings, which may in turn be reflected in differences in the regulatory framework concerning the collection and sharing of data. For example, developing countries may face more issues related to integrity of network data and misuse of personal information, due to a combination of weaker institutions and strong corporate power that may benefit businesses at the expense of

customers. The fact that businesses often operate in more than one context, makes things more complicated.

Second, the recognition that the increasing availability of online and digital data brings new ethical challenges due to the low cost and passiveness of data sharing in contrast to target active data collection with explicit consent and control of shared information. Progress has been made in raising consumer awareness and introducing tighter regulations, but we are far from achieving clarity and transparency. This is probably the area presenting the most ethical challenges for researchers nowadays, as it is also rapidly evolving.

Third, business imperatives driving a rapid expansion of digital services are providing firms with extensive surveillance capabilities both within the firm and beyond, with ethical implications for researchers engaging with such data that are yet to be fully understood.

A more fundamental dilemma that exists in social network research and which is further complicated in the setting under examination, is that the collection of individual data provides the means to reconstruct the collective behavior of the system. Therefore, an individual is not solely responsible for disclosing the global network structure, while the global structure is meaningless, or at least biased, without the contribution of all individuals. This shared importance blurs the notion of individuality and calls for reflections on the borders of individual responsibility and network data ownership not only for individuals, but also for organizations and researchers alike. This is especially the case in a business context, as the systematic analysis of business and management network research indicates that particular ethical concerns emerge at all stages. The ethical burden thus moves from the individual controlling which information to disclose, to the businesses controlling which information to collect from various stakeholders

and which to disclose to researchers, and on to the researchers controlling which methods to use for analysis and results to report. The entangled involvement of several stakeholders makes the design of a unified framework to avoid ethical dilemmas challenging.

References

- Abbott, A., 1988. Transcending general linear reality. *Sociological Theory*, 6(2), 169-186.
- Ahuja, G., Soda, G., Zaheer, A., 2012. The genesis and dynamics of organizational networks. *Organization Science* 23(2), 434–448.
- Aldred, R., 2008. Ethical and political issues in contemporary research relationships. *Sociology* 42(5), 887-903.
- Alvesson, M., 2003. Beyond neopositivists, romantics and localists: a reflexive approach to interviews in organizational research. *Academy of Management Review* 28(1),13-33.
- Badi, S., Wang, L., Pryke, S., 2017. Relationship marketing in Guanxi networks: A social network analysis study of Chinese construction small and medium-sized enterprises. *Industrial Marketing Management*, 60, 204-218.
- Baird, I.G., 2014. Principled engagement: obstacles and opportunities in an increasing consultancy dominated world. *ACME: An International Journal for Critical Geographies* 13(4), 497-507.
- Ball, K., 2010. Workplace surveillance: an overview. *Labor History* 51(1), 87-106.
- Beauchamp, T.L., Childress, J. F., 2001. *Principles of biomedical ethics*, 5th edn. Oxford University Press, Oxford.
- Bloom, P., 2019. *Monitored*. Pluto Press, London.
- Borgatti, S.P, Halgin, D.S., 2011. On network theory. *Organization Science*, 22(5), 1168–1181.

- Borgatti, S.P., Molina. J.-L., 2003. Ethical and strategic issues in organizational network research. *Journal of Applied Behavioral Science* 39(3), 337-349.
- Borgatti, S.P., Molina. J.-L., 2005. Toward ethical guidelines for network research in organizations. *Social Networks* 27(2), 107-117.
- Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M., 2011. The socialbot network: when bots socialize for fame and money. *Proceedings of the 27th annual computer security applications conference*, ACM, 93–102.
- Burt, R.S., Bian, Y., Opper, S., 2018. More or less guanxi: trust is 60% network context, 10% individual difference. *Social Networks* 54(1), 12-25.
- Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.F., Vespignani, A., 2010. Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One* 5, e11596.
- Chan, S., 2017. European court limits employers' right to monitor workers' email. *New York Times*, September 5. Available: <https://www.nytimes.com/2017/09/05/business/european-court-employers-workers-email.html>
- Chen, S., Lu, X., Liljeros, F., Jia, Z., Rocha, L.E.C., 2018. Indirect inference of sensitive variables with peer network survey. Under review.
- Conaldi, G., Lomi, A., Tonellato, M., 2012. Dynamic models of affiliation and the network structure of problem solving in an open source software project. *Organizational Research Methods* 15(3), 385-412.

- Covaleski, M., Dirsmith, M., Heian, J., Samuel, S., 1998. The calculated and the avowed: techniques of discipline and struggles over identity in big six public accounting firms. *Administrative Science Quarterly* 43(2), 293-327.
- Cronin, B., 2017. Limitations of Orbis relational data: UK director interlocks. Paper presented to the European Social Networks Conference, September 26, Mainz, Germany.
- Cronin, B., Popov, V., 2005. Director networks and UK firm performance. *International Journal of Knowledge, Culture and Change Management* 4, 1195-1205.
- Cross, R., Parker, A., 2004. *The hidden power of social networks: understanding how work really gets done in organizations*. Harvard Business Review Press, Cambridge, MA.
- D'Angelo, A., Ryan, L., 2019. The presentation of the networked self: Ethics and epistemology in social network analysis, *Social Networks* <https://doi.org/10.1016/j.socnet.2019.06.002>.
- De Montjoye, Y.A., Hidalgo, C.A., Verleysen, M., Blondel, V.D., 2013. Unique in the crowd: the privacy bounds of human mobility. *Scientific Reports* 3, 1376.
- Deville, P., Wang, D., Sinatra, R., Song, C., Blondel, V.D., Barabási, A.L., 2014. Career on the move: geography, stratification, and scientific impact. *Scientific Reports* 4, 4770.
- Economides, N., 1996. Network externalities, complementarities, and invitations to enter. *European Journal of Political Economy* 12(2), 211–233.
- Ellwardt, L., Steglich, C., Wittek, R., 2012. The co-evolution of gossip and friendship in workplace social networks. *Social Networks* 34(4), 623-633.

- Elovici, Y., Fire, M., Herzberg, A., Shulman, H., 2014. Ethical considerations when employing fake identities in online social networks for research. *Science and Engineering Ethics* 20(4), 1027-1043.
- Emirbayer, M., Goodwin, J., 1994. Network analysis, culture, and the problem of agency. *American Journal of Sociology* 99(6), 1411-1454.
- Ethics & Compliance Initiative 2016. *2016 global business ethics survey*. Ethics & Compliance Initiative, Arlington, VA. Available: <https://www.ethics.org/knowledge-center/2016-global-business-ethics-survey/>
- Faria, J.R., 2001. Rent seeking in academia: the consultancy disease. *The American Economist* 45(2), 69-74.
- Feighery, W.G., 2011. Consulting ethics. *Annals of Tourism Research* 38(3), 1031–50.
- Fishburne, P.M., 1980. Survey techniques for studying threatening topics: a case study on the use of heroin. PhD Thesis. New York University, New York.
- Fischer, C., 1982. What do we mean by ‘friend’? An inductive study. *Social Networks* 3(4), 287-306.
- Fitzgerald, L., Ferlie, E., Wood, M., Hawkins, C., 2002. Interlocking interactions, the diffusion of innovations in health care. *Human Relations* 55(12), 1429-1449.
- Fontes, L.A., 1998. Ethics in family violence research: cross-cultural issues. *Family Relations* 47(1), 53–61.

- Fortunato, S. and Hric, D., 2016. Community detection in networks: a user guide. *Physics Reports* 659, 1-44.
- Golle P., Partridge K., 2009. On the anonymity of home/work location pairs. In Tokuda H., Beigl M., Friday A., Brush A.J.B., Tobe Y. (Eds.), *Pervasive Computing*. Springer, Berlin, Heidelberg, pp. 390-397.
- Graham, J., Grewal, I., Lewis, J., 2007. *Ethics in social research: the views of research participants*. February. Government Social Research Unit, London.
- Greenwood, M.R., Holland, P., Choong, K., 2006. Re-evaluating drug testing: questions of moral and symbolic control. In Deckop, J., Giacalone, R., Jurkiewicz, C.L. (Eds.), *Human Resource Management Ethics*. Information Age, Greenwich, CT, pp. 161–180.
- Guillemin, M., Gillam, L., 2004. Ethics, reflexivity and “ethically important moments” in research. *Qualitative Inquiry* 10(2), 261–280.
- Hamilton, G.G. (ed.), 1996. *Asian business networks*. Berlin: Walter de Gruyter.
- Heemskerck, E.M., Young, K., Takes, F.W., Cronin, B., Garcia-Bernardo, J., Popov, V., Kindred, W., Winecoff, V.K., Henriksen, L.F., Laurin-Lamothe, A., 2018. Big corporate network data: problems, diagnostics, and fixes. *Global Networks* 18(1), 3-32.
- Hollenbeck, J.R., Jamieson, B.B., 2015. Human capital, social capital, and social network analysis: Implications for strategic human resource management. *Academy of Management Perspectives* 29(3), 370-385.

- Horak, S., 2018. Join in or opt out? A normative–ethical analysis of affective ties and networks in South Korea. *Journal of Business Ethics* 149(1), .207-220.
- Iphofen, R., 2009. *Ethical decision-making in social research: a practical guide*. Palgrave Macmillan, Houndmills, UK.
- Isaak, J., Hanna, M.J., 2018. User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer* 51 (8), 56–59. <https://doi.org/10.1109/MC.2018.3191268>.
- Kagan, D., Elovichi, Y., Fire, M., 2018. Generic anomalous vertices detection utilizing a link prediction algorithm. *Social Network Analysis and Mining* 8(1), 27.
- Karsai, M., Perra, N., Vespignani, A., 2014. Time varying networks and the weakness of strong ties. *Scientific Reports* 4, 4001.
- Kaur, J., Hoang, D.T., Sun, X., Possamai, L., JafariAsbagh, M., Patil, S., Menczer, F., 2012. Scholarometer: a social framework for analyzing impact across disciplines. *PloS One* 7, e43235.
- Kilduff, M., Oh, H., 2006. Deconstructing diffusion: An ethnostatistical examination of medical innovation network data reanalyses. *Organizational Research Methods*, 9(4), 432–455.
- Kim, T., McFee, E., Olguin, D.O., Waber, B., Pentland, A.S., 2012. Sociometric badges: using sensor technology to capture new forms of collaboration. *Journal of Organizational Behavior* 33(3), 412-427.

- Kitts, J.A., Lomi, A., Mascia, D., Pallotti, F., Quintane, E., 2017. Investigating the temporal dynamics of interorganizational exchange: Patient transfers among Italian hospitals. *American Journal of Sociology* 123(3), 850-910.
- Kogut B. (ed.), 2012. *The small worlds of corporate governance*. Cambridge, MA: MIT Press.
- Kurt, Y., Kurt, M., 2020. Social network analysis in international business research: An assessment of the current state of play and future research directions. *International Business Review* 29(2), 101633.
- Laffer, G., 2005. The critical failure of workplace ethics. In Budd, J.W., Scoville, J.G. (Eds.), *The ethics of human resources and industrial relations*. ILR Press, Champaign, IL, pp. 273–297.
- Langlois, A.J., 2011. Political research and human research ethics committees. *Australian Journal of Political Science* 46(1), 141-156.
- Laumann, E.O., Marsden, P.V., Prensky, D., 1983. The boundary specification problem in network analysis. In Burt, R.S., Minor, M.J. (Eds.), *Applied network analysis*, Sage, Beverly Hills, CA, pp. 18–34.
- Linhares, C.D.G., Travencolo, B.A.N., Paiva, J.G.S., Rocha, L.E.C., 2017. DyNetVis: A system for visualization of dynamic networks. *Proceedings of the 32nd ACM symposium on applied computing*. April 3-7, 2017. Marrakesh, Morocco.
- Lin, L.H., 2011. Cultural and organizational antecedents of guanxi: The Chinese cases. *Journal of Business Ethics* 99(3), 441-451.

- Liu, Y., Chen, L., Yuan, Y., Chen, J., 2012. A study of surnames in China through isonymy. *American Journal of Physical Anthropology* 148(3), 341-350.
- Lusher, D., Robins, G., Pattison, P.E., Lomi, A., 2012. “Trust me”: differences in expressed and perceived trust relations in an organization. *Social Networks*,34(4), 410-424.
- Mannheimer, S., Young, S.W.H., Rossmann, D., 2016. On the ethics of social network research in libraries. *Journal of Information, Communication, and Ethics in Society* 14(2), 139-151.
- Marineau, J.E., Labianca, G.J., Kane, G.C., 2016. Direct and indirect negative ties and individual performance. *Social Networks* 44, 238-252.
- Marsden, P.V., 1990. Network data and measurement. *Annual Review of Sociology*, 16, 435–463.
- Marsden, P.V., 2005. Recent developments in network measurement. In Carrington, P.J., Scott, J. Wasserman, S. (Eds.), *Models and methods in social network analysis*. Cambridge, UK: Cambridge University Press, pp. 8–30.
- Mehra, A., Dixon, A.L., Brass, D.J., Robertson, B., 2006. The social network ties of group leaders: Implications for group performance and leader reputation. *Organization Science*, 17(1), 64-79.
- Milojević, S., 2013. Accuracy of simple, initials-based methods for author name disambiguation. *Journal of Informetrics* 7(4), 767-773.
- Molina, J.L., Borgatti, S.P. 2019. Moral bureaucracies and social network research. *Social Networks*. doi:10.1016/j.socnet.2019.11.001

- Monaghan, S., Lavelle, J., Gunnigle, P., 2017. Mapping networks: exploring the utility of social network analysis in management research and practice. *Journal of Business Research*, 76(C), 136-144.
- Moreno, M.A., Goniu, N., Moreno, P.S., Diekema, D., 2013. Ethics of social media research: common concerns and practical considerations. *Cyberpsychology, Behavior, and Social Networking* 16(9), 708-713.
- Morton, A., 1999. Ethics in action research. *Systemic Practice and Action Research* 12(2), 219-222.
- Nespor, J., 2000. Anonymity and place in qualitative inquiry. *Qualitative Inquiry* 6(4), 546–569.
- Oliver, P., 2010. *The Student's Guide to Research Ethics*, 2nd edn. Open University Press, McGraw-Hill Education, Maidenhead, UK.
- Pallotti, F., Weldon, S., Lomi, A., 2020. Lost in translation: collecting and coding social interaction from audio-visual recordings. *Social Networks*, Special Issue on Social Network Data Collection. (Forthcoming).
- Quintane, E., Conaldi, G., Tonellato, M., Lomi, A., 2014. Modeling relational events: a case study on an open source software project. *Organizational Research Methods* 17(1), 23-50.
- Radicchi, F., Fortunato, S., Markines, B., Vespignani, A., 2009. Diffusion of scientific credits and the ranking of scientists. *Physical Review E* 80, 056103.
- Rocha, L.E.C., Liljeros, F., Holme, P., 2010. Information dynamics shape the sexual networks of internet-mediated prostitution. *PNAS* 107, 5706-5711.

- Rocha, L.E.C., Masuda, N., Holme, P., 2017. Sampling temporal networks: methods and biases. *Physical Review E* 96, 052302.
- Rossier, C., 2010. Measuring abortion with the anonymous third party reporting method. In Singh, S, Remez, S.L., Tartaglione, A. (Eds.), *Methodologies for estimating abortion incidence and abortion-related morbidity: a review*. New York: Guttmacher Institute; and Paris: International Union for the Scientific Study of Population. pp. 99-106.
- Schlæger, J., 2015. E-governance in Asia: an overview. In Schlæger, J. (Ed.). Conference proceedings. E-governance in Asia: strengthening transparency and accountability? 12-12 November. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2725860
- Sewell, G., 1998. The discipline of teams: the control of team-based industrial work through electronic and peer surveillance. *Administrative Science Quarterly* 43(2), 397–428.
- Shipilov, A., Gawer, A., 2020. Integrating research on inter-organizational networks and ecosystems. *Academy of Management Annals* 14(1), 92-121.
- Singer, E., 1978. Informed consent: consequences for response rate and response quality in social surveys. *American Sociological Review* 43(2), 144–162.
- Sykes, T.A., Venkatesh, V., 2017. Explaining post-implementation employee system use and job performance: impacts of the content and source of social network ties. *MIS Quarterly* 41(3), 917-936.
- Tamassia, R., 2013. *Handbook of graph drawing and visualization*. Chapman and Hall/CRC, Boca Raton, FL.

- Tilley, L., Woodthorpe, K., 2011. Is it the end for autonomy as we know it? A critical examination of the ethical principal of anonymity in the context of 21st century demands on the qualitative researcher. *Qualitative Research* 11(2), 197-212.
- TUC, 2018. *I'll be watching you: a report on workplace monitoring*. Trades Union Congress, London.
- Toole, J.L., de Montjoye, Y.A., González, M.C., Pentland, A.S., 2015. Modeling and understanding intrinsic characteristics of human mobility. In Gonçalves, B., Perra, N. (Eds.), *Social phenomena: from data to models*. Springer, Cham, pp. 15-35.
- Townsend, L., Wallace, C., 2016. *Social media research: a guide to ethics*. University of Aberdeen, Aberdeen, UK. Available: www.gla.ac.uk/media/media_487729_en.pdf.
- Tubaro, P. 2019. Whose results are these anyway? Reciprocity and the ethics of “giving back” after social network research, *Social Networks*
<https://doi.org/10.1016/j.socnet.2019.10.003>.
- Van Buren, H., 2001. If fairness is the problem, is consent the solution? Integrating ISCT and stakeholder theory. *Business Ethics Quarterly* 11(3), 481–499.
- Wade, D.T., 2007. Ethics of collecting and using healthcare data. *British Medical Journal* 334 (7608), 1330–1.
- Walford, G., 2005. Research ethical guidelines and anonymity. *International Journal of Research & Method in Education* 28(1), 83-93.

- Wasserman, S., Faust, K., 1994. *Social network analysis: methods and applications*. Cambridge University Press, Cambridge, UK.
- Wei, Y., Yildirim, P., Van den Bulte, C., Dellarocas, C., 2016. Credit scoring with social network data. *Marketing Science*, 35(2), 234-258.
- Wilkinson, D., Thelwall, M., 2011. Researching personal information on the public web: methods and ethics. *Social Science Computer Review* 29(4), 387-401.
- Windolf, P., 2002. *Corporate networks in Europe and the United States*. New York: Oxford University Press.
- Zaheer, A., Usai, A., 2004. The social network approach in strategy research: theoretical challenges and methodological issues. In Ketchen, D.J., Bergh, D.D. (Eds.), *Research methodology in strategy and management*. Amsterdam: Elsevier, pp. 67–86.
- Zang, H., Bolot, J., 2011. Anonymization of location data does not work: a large-scale measurement study. *Proceedings of the 17th annual international conference on mobile computing and networking* ACM, 145-156.
- Zhang, Q., Perra, N., Gonçalves, B., Ciulla, F., Vespignani, A., 2013. Characterizing scientific production and consumption in physics. *Scientific Reports* 3, 1640.
- Zuboff, S., 2019. *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. Public Affairs/Hachette, New York.

Table 1. Guidelines for considering ethical issues in business and management network research

Research stage / issue	Ethical Concerns	Mitigation
<i>Commissioning</i>		
Research questions	Gatekeeper interests can outweigh theoretical goals and interests of other stakeholders.	Care in project selection. Consider conflicts of interest.
Scoping	Pressure to stretch outcomes or ‘scope creep.’	Clearly specify realistic outcomes and limitations.
Transitory champions	Difficulty sustaining beneficial outcomes for all parties.	Specify all parties’ obligations.
Researcher dual identity	Dilemma in emphasis on research or client outcomes.	Specify mutually desired outcomes.
Management gatekeeping	Needs of some stakeholders not addressed.	Specify interests of all stakeholders.
Aggregated datasets	Data may be used for purposes other than providers intended.	Consider conditions under which datasets are compiled.
<i>Research design</i>		
Framing researcher relationship	Risk of treating data providers instrumentally.	Treat the relationship as researcher-participant.
Boundary definition	Risk of misinterpreting social situation of actors.	Careful consideration of boundaries and reporting of robustness checks and completeness statistics.
Spurious precision	Unwarranted interpretation may affect participants’ lives.	Report robustness checks on boundaries and missing data; report completeness statistics; develop and use scales for questionnaires.

Research stage / issue	Ethical Concerns	Mitigation
<i>Data collection</i>		
Quality of survey data	Task structure may generate unwarranted interpretation may affect participants' lives.	Careful construction of roster-based questionnaires, considering the task structure.
Identity ambiguity	Spurious identification of an individual with a particular network position.	Use data sources subjected to self-regulation or filtering; examine outliers in detail.
Observational techniques	Informed consent	Consider workplace surveillance context of data collection.
Collecting data on third parties	Collecting data on individuals without their consent.	Use precautionary approach. Care in avoiding opportunities for re-identification.
Use of data collected for one purpose for another	Informed consent for extended purpose. Business incentives to minimize compliance with data protection regulation.	Consider circumstances in which original data were collected.
Data ownership	Organizations have default ownership of data collected within it, despite any assurances given to participants.	In commissioning stage, clearly specify limits to use of data collected.
<i>Data analysis</i>		
Dataset combination	Provides information about participant characteristics without their informed consent; Increased risk of participant re-identification.	Restrict data to remote access and analysis or limited release.
Inference of missing data	Provides information about likely participant behaviour without their informed consent.	Consider value of research against high ethical risks.

Research stage / issue	Ethical Concerns	Mitigation
Choice of analytical technique	Inappropriate technique may misrepresent social position of individual.	Robustness checks around a range of techniques.
Culturally-biased theoretical constructs	Non-western sites categorized as in deficit of a western attribute.	Develop context-specific theoretical constructs.
<i>Reporting</i>		
Sociograms	Risks re-identification of individual participants and disclosure of valuable relationship information.	Report of metrics alongside abstract exemplar visualizations.
Choice of visualization technique	Risks distorted interpretation of social interactions of participants; Duty to communicate results clearly to participants.	Report algorithms used alongside robustness statistics and clear interpretation; Consider all stakeholders interests in reporting.
Disclose relevant findings to participants	Duty to report beneficial findings to participants.	Give early consideration to potential issues and specify during commissioning.
Ownership of report	Organizations have default ownership over what is reported, which may conflict with interests of other stakeholders.	Explore short embargos as a means to allow dissemination of results to other stakeholders.
Anonymity of research site.	Societal interest in specific activities of prominent firms and organizations.	Ask permission to identify research site, where organization is socially prominent.