Open Data for Open Innovation: managing skills and capabilities in SMEs

Franz Huber  
Seeburg Castle University  
Business Administration  
fra.huber@gmail.com

Francesco Rentocchini  
University of Milan  
Department of economics, management and quantitative methods  
francesco.rentocchini@unimi.it

Thomas Wainwright  
Royal Holloway, University of London  
Tom.Wainwright@rhul.ac.uk

Abstract

Open Data (OD), data published online and in machine-readable format for everybody to use and republish without financial costs or restrictions, have been encouraged by governments because of their potential to fuel digital innovation. Utilizing OD could enrich open innovation opportunities for small and medium-sized enterprises (SMEs) to overcome their resource-constraints. Yet, the extent to which these opportunities have been exploited and how the processes of OD-based open innovation operate have not been examined by academic research. Furthermore, there is a lack of research on the internal capabilities and human capital that is required for resource-constrained SMEs to benefit from open innovation. Drawing upon 30 semi-structured interviews with UK-based organisations, which publish or consume OD, this paper aims to address these gaps. The results specify the inbound and coupled open innovation mechanisms of how SMEs can benefit from OD. The paper argues that OD use is not “free”, but involves considerable costs and barriers for SMEs, including difficulties in assessing the usefulness of OD, integrating and analysing data, or legal risks and risks of imitation. Finally, the paper qualifies the individual skills and organizational competencies that SMEs require to benefit from OD-based open innovation. This ranges from the technical aspects related to data science, to entrepreneurial abilities to identify OD opportunities when developing OD-based business models and to engage with external OD publishers and users. Overall, the results suggest that the resource-constrained nature of SMEs makes it difficult for them to fully exploit the innovation potential of OD, and SMEs require support from public policy and other actors in the OD innovation ecology.
Abstract

Open Data (OD), data published online and in machine-readable format for everybody to use and republish without financial costs or restrictions, have been encouraged by governments because of their potential to fuel digital innovation. Utilizing OD could enrich open innovation opportunities for small and medium-sized enterprises (SMEs) to overcome their resource-constraints. Yet, the extent to which these opportunities have been exploited and how the processes of OD-based open innovation operate have not been examined by academic research. Furthermore, there is a lack of research on the internal capabilities and human capital that is required for resource-constrained SMEs to benefit from open innovation. Drawing upon 30 semi-structured interviews with UK-based organisations, which publish or consume OD, this paper aims to address these gaps. The results specify the inbound and coupled open innovation mechanisms of how SMEs can benefit from OD. The paper argues that OD use is not ‘free’, but involves considerable costs and barriers for SMEs, including difficulties in assessing the usefulness of OD, integrating and analysing data, or legal risks and risks of imitation. Finally, the paper qualifies the individual skills and organizational competencies that SMEs require to benefit from OD-based open innovation. This ranges from the technical aspects related to data science, to entrepreneurial abilities to identify OD opportunities when developing OD-based business models and to engage with external OD publishers and users. Overall, the results suggest that the resource-constrained nature of SMEs makes it difficult for them to fully exploit the innovation potential of OD, and SMEs require support from public policy and other actors in the OD innovation ecology.

1. Introduction

Policymakers and researchers have highlighted the important role of innovative small and medium-sized enterprises (SMEs) in job creation and economic growth (Wolf & Pett, 2006). More recently, digitization and Internet-based opportunities for information sharing have transformed the potential for open innovation (OI) in SMEs (Bogers et al. 2017; Dodgson et al., 2006; Huber, 2013; Whelan et al., 2010). Data in particular can be viewed as the 'raw material' for the digital economy. This paper focuses on new developments in Open Data (OD), data that are published by organisations online and in machine-readable format, for everybody to use and republish without financial costs or restrictions (Open Data Institute 2015).
Researchers have argued that OD offers new opportunities for organisations, particularly SMEs, to develop new value-added applications or services as it is freely accessible (Chan, 2013; Janssen, 2011; Lee et al., 2014). Similarly, scholars have argued that OI can assist SMEs in overcoming resource and capability constraints (Eftekhari and Bogers, 2016), indicating that there is clear potential for OD based OI in SMEs. Despite this, research by Almirall (2015) and Lee et al. (2014), has suggested that OD has so far been unsuccessful in meeting its full potential, with organizations failing to commercialize OD into new digital products and/or services. This raises the question: do SMEs struggle with OD OI? There is limited research on the processes of OI in SMEs, and, specifically, a lack of research on the internal competencies and human capital that they require to exploit the potential of OI (Wynarczyk et al, 2013). These puzzles raise questions which make OD based OI in SMEs a fruitful area of academic research. Using the context of OD, we examine the understudied processes of SME OI. This paper also identifies the advantages and difficulties in successfully commercializing OD in the digital economy.

Our paper aims to explore (i) the barriers to OD use; (ii) the specific digital competencies and skills that shape OD OI activities in SMEs and (iii) to what extent SMEs have benefited from the OD based opportunities. Based on 30 semi-structured interviews with organisations in the UK which publish and/or consume OD, our paper provides novel insights into the difficulties and benefits of pursuing OD-based OI. Our conceptual model reveals various types of costs and barriers in OD use, and we identify the critical digital skills and organisational competencies that are required for utilising OD in OI activities, to overcome resource constraints.

The remainder of this paper is structured as follows: section 2 will outline the context of OD and the literature on OI for SMEs. Section 3 describes the methodology. Section 4 examines types of barriers and costs of OD, the unique human capital and organizational capabilities that support SMEs in capturing value from OD based OI, and the advantages of OD to SMEs. Section 6 will conclude the paper.

2. Open innovation and open data for SMEs

2.1 Open data and the digital age

Data is the 'raw material' of the digital age with innovators seeking new sources of data to create new value-added products and services through OI (cf. Bogers et al. 2017). While researchers have increasingly drawn attention to the importance of 'big data' (Chen and Zhang, 2014) and 'linked data' (Wood, 2010) in the digital economy, the recent phenomenon of OD remains comparatively understudied. The OD Movement
began in 2009, when the US government issued a Memorandum on Transparency and Open Government, which aimed to make government data more public (Lee et al. 2014). Since 2009, European and emerging economy governments have also begun to publish OD, in addition to private sector organizations, data that can be used in SME OI. OD is digital and available in a machine readable format, which enables the data to be used directly in innovative and value-added applications or services (Chan, 2013; Janssen, 2011). OD often comprises of data on activities such as transportation, public and private sector organisations, location based services, property, education, health and meteorological activities, and is viewed by policy makers as an important resource for the growing digital economy.

Although there are some similarities, OD innovation is different to open source software innovation (Lindman and Lyman, 2014), another important part of the digital economy. Piva et al. (2012) argued that open source software can help SMEs overcoming internal resource constraints and can provide access to external competences and complementary assets. In contrast, OD is about data as a ‘raw material’ or asset for a wide range of online applications, whereas open source is centred on licenses for applications as a compiled source code. Because of these differences, OD innovation requires different knowledge and competencies, compared to open source innovation.

OD has attracted increasing interest from policymakers and practitioners, but recent academic studies are sporadic. On the one hand, given the technical roots of OD in webscience, this discipline has examined technical processes and techniques including the, semantic web, ontologies and linked data, to advance the practical applications of OD (Jain et al. 2010; Missier et al. 2010; Oren et al. 2008). On the other hand, information management studies have examined OD’s role in open government policies to enhance transparency, democracy and civic engagement (Janssen et al. 2012; Bates 2012; 2013). Despite this, there is limited research on OD in management and innovation studies (for exceptions see Beronne et al. 2016; Perkmann and Schildt, 2014) raising questions around how OD can be used by SMEs in OI.

### 2.2 Open innovation for resource-constrained SMEs

It has often been argued that SMEs face limitations in innovation due to a lack of slack resources and difficulties in forming external partnerships (Hewitt-Dundas, 2006), which can impede the identification of new opportunities and external knowledge (Bianchi et al., 2010; Prajogo & McDermott, 2014; Maes & Sels, 2014). It has been suggested that OI mechanisms, as “purposefully managed knowledge flows across organizational
boundaries” (Chesborough and Bogers, 2014, p.17), can address these challenges (Eftekhari and Bogers, 2016). Knowledge sharing occurs through inbound and outbound OI mechanisms of a pecuniary or non-pecuniary nature (Dahlander and Gann, 2010; Freel and Robson, 2016). Inbound OI refers to outside-in processes embracing external inputs, whereas outbound OI (inside-out), involves the opening of underutilized or unused assets to actors outside of the organization through processes of revealing, sales or licensing. Coupled OI refers to the combined use of inbound and outbound flows (Dahlander and Gann 2010).

Previous literature has shown that OI increases new product development (Dahlander and Gann, 2010; Laursen and Salter, 2006), revenue growth (Chesbrough and Crowther, 2006) and financial performance (Rothaermel and Alexandre, 2009). However, this evidence has been mainly confined to large companies, while SMEs have been comparatively overlooked (Spithoven et al., 2013). Research on SMEs in six European countries by Lasagni (2012) show that SMEs with stronger supplier, user and customer relationships, and those with research institute linkages, have a higher innovation performance. Also, recent work has found that network formation with external partners can help overcome the liability of smallness (Colombo et al, 2012; Freeman et al., 1983; Gassmann et al, 2010), while enabling business and technological knowledge to be sourced (Huber, 2013). Furthermore, the embeddedness of SMEs in broader relational and institutional contexts matters for OI, as discussed by literature on network structures (Slotte-Kock and Coviello, 2010) and broader ecosystems (Nambisan and Baron, 2013).

Studies have argued that benefiting from inbound OI does not happen automatically, but requires absorptive capacity, the ability to value, assimilate and to apply new external knowledge or assets (Cohen and Levinthal, 1990; Xia and Roper, 2016). SMEs, in contrast to larger firms, tend to have less resource for utilizing external relationships (Colombo et al, 2012) and show a lower tendency to collaborate with external partners (Ebersberger et al., 2012), although their flexibility and lack of core rigidities, can benefit them in utilizing new external opportunities (Colombo et al, 2012; Spithoven et al., 2013). SMEs require internal management and organizational capabilities to assimilate and apply external knowledge in inbound innovation activities (Brunswicker et al., 2015; Robertson et al, 2012), and a lack of multidisciplinary competencies (Bianchi et al, 2010) and less structured approach to innovation management (De Toni and Nassimbeni, 2003) may restrict them. For a recent review on the determinants, motivations and barriers of OI with relevance for SMEs see Verbano et al. (2015), and Parida et al (2015). In contrast to larger corporations, SMEs have received limited attention by OI scholars (Chesbrough and Bogers, 2014; van de Vrande et al., 2009), which may explain
the puzzle surrounding potential OI SME benefits and opportunities. On the one hand, OI appears to have obvious advantages for resource-constrained SMEs, providing them with access to new knowledge. On the other hand, the ability of organizations to capture value from OI, can also be potentially constrained by SMEs’ limited resources and absorptive capacity. This raises questions over which of the above arguments holds. Initial research indicates that this puzzle is also relevant to the case of OD. Key protagonists of the OD Movement have argued that data can be used to create new digital innovations, although widespread success has so far been illusive (Lee et al. 2014), despite the notion that access to free data creates new opportunities for OI. For example, Almirall (2015) argues that the best case for OD use in the US is data.gov, with 400,000 datasets, published by 200 government agencies, but that only 150 apps emerged from data on the portal, and only 24% of those have more than 10,000 downloads. This highlights the difficulties in commercializing OD, and that greater data availability and OI does not necessarily support the digital economy.

Researchers still lack a sufficient understanding of the exact mechanisms of how OI benefits SMEs (Eftekhari and Bogers, 2016; Verbando et al., 2015), with Brunswicker et al (2015), stressing how external knowledge sourcing in particular remains a poorly understood aspect of OI. OI should not be assumed to necessarily lead to improved performance, because it is context-dependent (Dahlander and Gann, 2010, Gassmann, 2006; Laursen and Salter, 2006; Lazzarotti and Manzini, 2009; Verbano et al, 2015) and further research is needed to gain a better understanding of the unique internal capabilities required for SMEs to successfully integrate and exploit external knowledge sources (Brunswicker et al, 2015).

3. Methodology

This paper draws upon an interpretive qualitative approach (Gephart 2004) to obtain real-time and retrospective accounts how OD based OI takes place, particularly the advantages and constraints of using OD and the skills and competencies required. The data consists of 30 semi-structured interviews with experts working in UK organisations adopting and OD-based OI strategy, complemented by online information and documents of the respective organisations.

The UK is central to the OD Movement, as it was one of the first countries to introduce legislation requiring government departments to publish OD in 2012 (BIS 2014). Furthermore, the activities of key OD protagonists including Sir Tim Berners-Lee and Sir Nigel Shadbolt, are based in the UK. This approach enables us to address the gaps
identified in Section 2, by providing insight into the different processes involved in publishing and consuming OD through OI processes, based on variations in organisation type. Understanding how OD is published by public and corporate organisations is important to determine how OD is positioned for its inbound use by SMEs. We deliberately chose to target a variety of organisations, including SMEs, public sector OD publishers and larger corporations to gain a more holistic picture of the OD ecology and to provide insight into the research question from different angles. Keeping a broad view at this stage seems also relevant in relation to the recent interest developed within the OI literature on the important role of platforms and ecosystems for OI (West et al., 2014). To select relevant cases, the researchers used the ODI’s membership list and examples from data.gov, along with snowballing, to access respondents in the OD developer community. In total, 37 OD organisations were contacted, 30 of which responded, and we could not identify clear patterns for non-response. The characteristics of the organisations interviewed are shown in Appendix 1. Following the established ‘key informant interviews’ approach (Squire et al, 2009), interviewees were key decision makers on OD in the respective organisation, and the job positions included chief information officers, heads of data and statistics, owner-managers, and product/innovation managers (see Appendix 1).

A semi-structured interview guide was used to ensure consistency and that similar issues were examined in all cases, with sufficient flexibility and room to capture issues that interviewees thought important to their organisation (cf. Biniari 2012). Interviews lasted between 45 minutes and two hours, were recorded and later transcribed. Face to face interviews were conducted by all three authors in the UK, between 2014-2015 to assist in the contextualisation of the information provided, although some organisations insisted on Skype interviews because of flexible working routines (see Appendix 1). The interviews examined the experiences for using inbound and outbound OI, with a focus on the capabilities and skills needed to absorb OD within the organisation. Credibility probes were used in the interviews to undermine strategic impression management (Rubin and Rubin, 2011). In addition, organisational websites, API blogs, company reports and documents were examined when available. Respondent data was anonymised to protect the identities of the interviewees and their organisations.

The analysis used a multi-step iterative process, drawing upon theoretical issues identified earlier in the OI literature (Section 2) and inductive reasoning to shape the conceptual development (Miles et al., 2013). Inspired by the Gioia methodology (Gioia et al, 2012), we aimed to develop a data structure by first identifying 1st order categories (interviewee-centric terms). Then we explored linkages and patterns within 1st order categories and developed 2nd order themes (researcher-centric themes). Afterwards, we
synthesised the 2nd order themes into aggregate dimensions. Whilst the interview data constituted the core for these steps, we also utilised information on organisational websites, API blogs, company reports and other documents, if available, to verify or enrich the arguments. First, we adopted this procedure for examining the barriers and costs of using Open Data (Figure 1). Second, we explored what is required for SMEs to utilise OD for open innovation (Figure 2). Third, we investigated how SMEs can benefit from OD (Figure 3). These three data structures are synthesised in the conceptual framework (Figure 4), which provides an overview of the contributions of the paper on factors that influence the process of OD OI in SMEs.

Our analysis used critical verification techniques to optimise validity in the interviews (Morse et al., 2008), which involved double-checking coded themes and their interpretation, whilst undertaking a critical interpretation of the transcripts, to avoid biased memorising. To optimise validity, we applied data triangulation (utilising online information and documents in addition to the interview data) and investigator triangulation (three researchers collected data and two researchers independently coded 1st order concepts and 2nd order themes) - (Denzin, 1978). Whilst there are challenges in developing robust generalisations from qualitative studies, this study proposes a conceptual framework on the qualitative types of enablers and barriers to OD based OI and on the skills and capabilities needed to successfully develop new digital applications and services.

4. Results

4.1. Open Data is not ‘free’: barriers and costs of using Open Data

The results reveal a range of barriers and costs to SMEs using OD. Figure 1 provides an overview and illustrative quotes of the various reasons leading to second-order themes and aggregate dimensions.

First, given the growing number of OD sources published, there is a need to be able to horizon scan and find opportunities for using relevant OD. Several SMEs in the sample felt that they cannot take full advantage of OD because they do not have the resources to look for useful OD in a crowded ecosystem of growing datasets. Second, as
mentioned in the previous section, much of the value added to OD comes from merging and linking it with other open or proprietary sources of data. Integrating various OD sources often generates considerable work, especially when it concerns different types of publishers, perhaps even in various regions or countries. OD publishers often release data following different formats and standards, while internationally scoped OD uses different languages. This generates work to prepare the OD for analysis and can necessitate building relationships with OD publishers. Third, even if OD is suitable for immediate analysis, it usually requires data science skills. Many SMEs struggle to obtain those skills, given the limited supply of data scientists on the labour market, combined with SMEs’ limited financial resources. Fourth, several SMEs stressed that the uncertainty of the future availability or quality of OD represents a considerable disincentive for investing in developing OD centred apps or services. SMEs can become vulnerable if their business model is heavily dependent on future publishing and updating of OD, and there is a risk of its discontinuation or withdrawal. Fifth, respondents stressed legal and associated reputational risks stemming from uncertainties regarding the responsibilities for potentially using incorrect data in their digital products or services. Clarifying these issues would often require considerable legal expertise that is usually not readily available in SMEs. Finally, a sixth reason is the risk of competitors imitating a SMEs app or service, which results from the public nature of OD. Given the above-mentioned costs faced when utilizing OD, managing this risk of imitation is a key challenge. Overall, the results illustrate that despite of their open and free nature, utilizing OD is not free of costs and risks.

[FIGURE 1 ABOUT HERE]

4.2. Capturing value from open data: human capital and organizational capabilities

Identifying the costs and barriers to OD use, raises the critical question of which skills and capabilities are required for SMEs to be able to overcome these barriers and utilize OD effectively. As illustrated in Figure 2, two distinctions proved useful: the individual versus organizational scale—corresponding to human capital based skills versus organizational capabilities—and business management related entrepreneurial versus technical/professional skills/capabilities.

[FIGURE 2 ABOUT HERE]

4.2.1. Entrepreneurial skills and competencies
Whilst the paper identified several technical challenges of using OD above, several respondents argued that the most critical challenge concerns the business mindset of venture teams in appreciating the opportunities of OD based OI (see the illustrative first quote in Figure 2). This requires an ‘open’ approach in using OD, as well as skills in developing a robust business case, potentially including a bricolage of open and closed data and the ability to evaluate the value added for its users. Within this context, SMEs need to find ways of dealing with the potential threat of imitation by competitors, and relationships with clients and the ability of continuous innovation can be critical. A viable business often requires skills in managing relationships with external partners or developer communities. On the organizational scale, this often has to be in an institutionalized form, with regular engagement with OD publishers and data consumers. On an organizational level, there can also be resistance to opening data, with key evangelists needing to convince the management team who are concerned with risks, about the potential benefits.

An organizational culture which embraces and supports OD based innovation can be critical, although the most frequently mentioned capability for utilizing OD is the ability to hire, retain and integrate a mix of talent, which bridges entrepreneurial management competencies, with the technical or professional skills, outlined in the next section.

### 4.2.2. Technical/professional skills and competencies

In general, many respondents highlighted that hiring suitable data scientists is a considerable challenge, even more so for people that can bridge data science and digital marketing. For SMEs operating internationally, international language skills and geographical knowledge are also critical factors:

“The biggest challenge that we face is we are trying to do this for many markets, so you need people… I want someone who can think in 5 different countries… And so a big challenge that we have is getting people the right mix of languages” (Case 16)

The most difficult challenge is to have a mix of technical skills and user-centred innovation skills in the SME:

“You do need that data scientist; you do need the people to pick up the trends and drive the correlations to create the models, but you also need someone that’s on the ground that can say, ‘Well actually if I knew this, I can then take advantage of it.’…” (Case 30)
Bridging technical data skills with value creation for consumers in the respective domain was identified as the most critical, and most challenging, requirement for OD-based innovation.

Finally, as already mentioned in section 4.2., legal knowledge, as well as knowledge in licensing/creative commons is important to deal with legal and reputational uncertainties.

Overall, acquiring these technical skills or capabilities required for absorptive capacity in the digital age (Roberts et al., 2012) is non-trivial and compared to larger companies, smaller SMEs are likely to have less capacity to develop them.

4.3. How OD can support SME Open Innovation

The advantages of using OD, as highlighted by the interviewees, mainly refer to inbound OI and coupled OI (combined use of inbound and outbound flows). Figure 3 provides an overview of key activities where OD can support SME innovation, including illustrative quotes.

The most obvious advantage is the inbound OI mechanism, where OD being available for free, removes the need to pay for proprietary data. All SMEs in our sample mentioned that this is beneficial given their limited resources. As a second inbound OI mechanism, several respondents stressed that OD enabled access to novel data that would not be available otherwise, examples including live public transport data on train and bus services, property prices, and public health data. A third type of inbound OI, potentially integrating the above-mentioned types, is centred on combining external OD with internal proprietary data without making them public again. OD provides an opportunity for recombining existing data, and integrating it with proprietary internal data can enable the development of specific digital apps or services that are difficult to imitate by competitors.

A more complex benefit of OD is achieved through a coupled OI mechanism. SMEs source external OD, then creatively link it to their own internal data or other proprietary sources before publishing this new dataset as OD. The new OD attracts attention from external users, who will then utilize the OD in their own apps and innovations. In some cases, external developers would create an affiliate app or website with the main intention of feeding more traffic to the original OD publishing SME. This can enable the original SME to upscale its activities and income through coupled OI. In one case, this was used as part of a low-cost internationalization process, where local overseas developers familiar with a geographical market would create apps for free using external OD, enabling customers to buy from the original OD publishing SME, but using a different payment system and language. Affiliate SMEs are paid a small fee for sending
web traffic to OD SME publishers. Through coupled OI, the original OD publisher benefits from access to a larger market, while the inbound affiliate received referral and traffic fees, through sending traffic to the OD publishing SME. Overall, OD-based coupled OI can enable SMEs to attract attention and to upscale with low costs.

[FIGURE 3 ABOUT HERE]

5. Conclusions

This paper has provided new insight into the potential of one particular movement in the digital economy: OD-based OI in SMEs. Whilst the free and public nature of OD provide potential opportunities for SMEs to overcome resource-constraints, studies have indicated that SMEs struggle to successfully utilize OD in innovation, creating a puzzle. Within the OI literature more broadly, SME OI remains understudied (Wynarczyk et al., 2013), when compared to the volume of studies that consider OI in well-resourced corporations. There is also a particular lacuna of understanding concerning the skills and competencies needed to absorb and modify external assets in SMEs (Brunswicker et al, 2015). To address these gaps and solve the puzzle, our work identifies what prevents the adoption of OI practices in the OD realm (OD inhibitors), it then turns to highlight the individual and organizational factors which allow SMEs to overcome these barriers (OD enablers). Lastly, it shows how OD contributes to inbound and coupled OI. Overall, the paper provides new insights into how SMEs can benefit from the release of OD as part of OI mechanisms. Figure 4 provides an overview of the conceptual model derived from the empirical analysis.

[FIGURE 4 ABOUT HERE]

The first contribution revealed the specific inhibiting factors as to why utilizing OD is not ‘free’, but involves costs and barriers as well as future uncertainty and risk. We examined how scoping appropriate and useful OD was complex, as was obtaining the resources, skills and capabilities to use OD in new digital app and service design, while ensuring that the OD was sustainable and would be available in the future. Managing the risk of imitation by competitors was also a concern for SMEs, as the OD is freely available to all, in addition to balancing the legal and reputational risk of using external OD and in publishing their own OD. The results substantiate why digital innovation with freely available data requires specific aspects of absorptive capacity (Roberts et al., 2012).

The second contribution is to qualify the enabling factors of which individual skills and organizational capabilities are required for benefiting from OD. The results showed that on the one hand this concerns technical/professional skills related to data science, data-driven marketing and legal knowledge that are in short supply. On the other hand,
entrepreneurial skills in appreciating the potential of OD and in developing a robust business case could be identified as critical skills that tend to be underdeveloped. On an organizational level, benefiting from OD requires an organizational culture that embraces OD and institutionalized engagement with OD publishers and consumers, as well as the ability to hire, retain and integrate a mix of technical and business talent. This result resonates well with the recent findings showing that professional and innovation training for employees improves attitudes towards OI by effectively reducing Not-Invented-Here and Not-Shared-Here syndromes (De Arujo Burcharth et al., 2014). In general, the organizational capability of continuous innovation is critical for those OD centred SMEs that face the potential threat of imitation. Overall, these substantial requirements for benefiting from OD identified by this study suggest that the well-known problem of resource constraints in SMEs bites again (Hewitt-Dundas, 2006), as SMEs tend to struggle to develop internal organizational capabilities to utilize external knowledge (Brunswicker et al, 2015; Robertson et al., 2012) and are often restricted by a lack of multidisciplinary competencies (Bianchi et al., 2010). Developing those skills and capabilities is non-trivial, especially for micro and small firms.

The third contribution of the paper was to specify how OD opens up new opportunities for SMEs through OI, particularly in reducing proprietary data costs and making available entirely new datasets as inbound OI, which enable the development of novel digital apps and services. SMEs are also able to expand their activities through coupled OI, enabling resource constrained SMEs to grow their income and scale through affiliated SMEs which utilize OD. In general, the combination of external OD with other OD or proprietary data enable the development of innovative apps and services. This brings to the forefront the importance for SMEs to develop strategies to simultaneously enable and filter external interactions: the development of platforms and ecosystems as well as the fine-tuning of appropriate business models to reap benefits will play a central role for the future survival of OD-based OI SMEs in the digital economy (Piller and West, 2014; West and Bogers, 2014; Altman and Tushman, 2017).

The ability to deal with and utilize data, whether open or proprietary, is a key challenge for innovation in the evolving digital economy, and further study is needed to examine resource constrained strategies for integrating external data in more detail, and to examine how coping strategies and additional support mechanisms could be developed to assist OI in SMEs. Furthermore, future research should examine the roles various types of organisations (public and private organisations of different sizes) can play in the emerging OD innovation ecology (van der Borgh et al., 2012). Developing more detailed insight into these areas will enable practitioners and policy developers to address the barriers to SMEs in the digital economy. The results of this paper support the claims of
key proponents of the OD movement (e.g. Cabinet Office, 2012; Open Data Institute, 2015) that OD provides an opportunity space for private sector innovation and for entrepreneurial SMEs. Yet, the results highlight and specify why utilizing OD is not ‘free’ but requires considerable skills and competencies. Addressing how SMEs can be supported best, and how the players in the OD innovation ecology can interact and cooperate to unleash the potential of OD for SME innovation represents a key challenge for policy developers, academics and private sector practitioners.

Bibliography


Piller, F., & West, J. (2014). Firms, users, and innovation. H. Chesbrough, W. Vanhaverbeke and J. West (Eds.), New frontiers in open innovation, Ch. 2, Oxford University Press.


Figures
Figure 1: Types of barriers and costs of open data

First-order categories

(illustrative quotes)

"...last time I checked, there was something like 8,000 or 9,000 datasets available on data.gov, so it's pretty difficult to stay on top of new data sources [...] we just don't have the time and resources." (Case 6)

"...it's very much a numbers game. So they [OD publishers] can stand up and say 'Look, we've got 45,000 datasets, isn't that great. But it's very much a case of push it out, push it out, push it out. But then, that's what also adds the value to the market, because other companies wouldn't have these skills and expertise to keep following these things." (Case 7)

"Our main problem is that whilst open data is our friend, it is also a double edged sword, because at the moment we're working across 45 countries, so that you can imagine the amount of different municipalities, cities, organisations in those countries, and we're having to put all the data into one standard format that we can then use internally. So that in itself presents a big problem..." (Case 16)

"[T]o create something of interest, it takes work, effort and time. Are they going to have three or four developers sat there monitoring new data releases, collecting the data, explaining it to other people? Probably not, and if they do, it would only be the very top group of companies that do that, so there's still quite a vast market underserved that..." (Case 25)

"[They've] thrown all this data out there, I feel that we haven't got the resources to work with that data to understand it." (Case 15)

"There aren't any data scientists. There aren't. It's so hard to find...Yeah, but there just aren't any. It's only recently that data has become sexy. [...] No, it's a vastly under resourced area, in my experience, they're very hard to get." (Case 21)

"I think there's really one major risk, which is continuity of supply. So if we based the business on using open data, which we are obviously not paying for, then it's very difficult for us to have any lever to ensure that that data continues to be available. So political changes, or if an organisation changes its mind about publishing, we are likely to be a bit stuck. [...] It would probably kill any product using that data." (Case 12)

"Well at that time I was doing some work with [x] and we'd put a lot of effort into preparing for the next flow survey [Open Data]. We had developed a prototype [app], everyone loved it, and then it [Open Data] got pulled.” (Case 5)

"...published [open] data, and if it's wrong, can somebody sue me? Can someone bring a law suit against me for having inaccurate data like this...There's a whole series of questions like that which get lawyers super-excited...And then the other perspective, confidentiality and security, national security, public data, what if I advertise the weaknesses of the electrical grid, is that open data or am I asking for an attack? I have spent a lot of time on that..." (Case 23)

"The main benefit is the main downside as well, because everyone has access to it [...] We started developing the app and we came up with it, we found that there were another 4/5 little companies coming up with solutions. The second is, it drives the innovation, because you really have to be much better than all the others." (Case 9)

"anyone can actually use our data now... and the thing I'm most wary of is, is a rival, and he just copies all our data and tries to do the same thing using all our hard work." (Case 14)
Figure 2: Overview of human capital and organisational capabilities for utilising Open Data

First-order categories (Illustrative quotes)

- “[...] it’s mentally how you think about your business, it’s to commit to doing things openly, it’s technically and physically very easy, […] To commit to being open as a business is a much harder thing, because it requires you to do loads and loads of work” (Case 29)
- “I mean it’s open source, so they can do that [stealing ideas] if they wish to. I guess there are two protections that we have in reality. Number 1 protection is we have relationships with our clients, so once we have that relationship it’s difficult for someone else to copy our product and hence try and steal that relationship” (Case 22)
- “One of the hardest jobs we had to recruit was a data analyst with marketing focus” (Case 2)
- “So you need that mix of data scientist and practitioner. So then you can then say “This is what I want” [...] We’ve reached the point where look [...] just stop proving the technology, let’s start doing something with this information... it has to be demand driven” (Case 86)
- “There are definitely licensing agreements around that and we are still working them out. We have a contract with Y and they have given us actual backing [access to OD and additional proprietary data], but again it seems to me they are not even 100% clear on what can be used and what can’t be.” (Case 17)
- “Ultimately end up, with one or two enthusiasts within the organisations that really want to do it (using OD), but they still have to convince their own management that it’s a good idea. So yes, there are often various objections about things like “Well, so it’s going to cost us money; what’s in it for us?” is one obvious thing. Or problems about they’re worried of the data quality or they’re worried about people misunderstanding it or misusing it, these kind of things” (Case 19)
- “We also interact with businesses using open data, particularly through the OD1, which is obviously quite an essential point for these businesses. So that enables us, I guess, to find out who is publishing interesting data and to share ideas about how to use it... when we are setting something up with an open data source, we will informally probably discuss with the publisher, and also with the client” (Case 14)
- “We are trying to balance the consumer side to being able to think like the consumer, which is kind of lets say more qualitative. On the other hand, we need to understand all the data which is an important quantitative type skill. So in the team we need both the splits” (Case 16).
- “we continue to innovate, so somebody can copy what we are doing now, but then they have to continue to copy any new innovations.” (Case 12)

Second-order themes

- Ability to develop a robust business case with OD
- Relationship management skills
- Skills in bridging data science and marketing
- Knowledge on legal aspects of data
- Organisational culture for open data OD
- Ability to engage with OD publishers and consumers
- Ability to hire, retain and integrate a mix of talent
- Capability for continuous innovation

Aggregate dimensions

- Individual entrepreneurial skills
- Individual technical/professional skills
- Organisational entrepreneurial capabilities
Figure 3: Types of open data advantages

First-order categories (Illustrative quotes)

"so we couldn’t do it without open data” (Case 30)

"But generally speaking, I think if there hadn’t been at least a lot of open data datasets available, we probably wouldn’t have had the product in the first place” (Case 23)

"[We] used to pay for these mountains of data we are getting now for free, they will simply add more value in different areas of the business.” (Case 9)

"It’s free! But I mean, then it’s – yeah. It’s free. So it means that as an organisation or a business, you can get access to information that hitherto would have cost a hell of a lot of money” (Case 5)

"It’s sort of secondary and tertiary uses where individuals and private companies like ours can take this [OD] and mash it up and create some new weird and wonderful tool for the private market. [We] use open data as a foundation, but the value that people get from that data is…combining that with other types of proprietary data (Case 8)

"Look, this is what other people are doing, and on their own [OD] has mediocre value, but when meshed together [with other data], it has an exponential value.” (Case 2)

"We have our published [OD] channels... start-ups in foreign countries [...] particularly in areas like India... it was an area for us to go into their market without heavy investment, because all we had to do was release our data; someone else does all the work, and turns it into a development.” (Case 2)

"Also because it benefits our business so that the way we make money is we make our database available and people click on properties, and we then send those leads on to our customers. So the more people I can get seeing my database the better... if someone can generate a significant amount of traffic for us, we’ll gladly pay them for that” (Case 8)
Figure 4: Overall conceptual model: from OD to OI

OD publishers

Inhibitors for OI
- OD barriers and costs:
  - Difficulties in identifying OD sources and usefulness
  - Time and skill resource needed to integrate OD
  - Shortage of data scientists who can utilise OD
- Future uncertainty and risk:
  - Uncertainty in future availability of OD
  - Legal and reputational risks
  - Risk of imitation

Enablers for OI
- Individual skills:
  - Ability to develop a robust business case with OD
  - Relationship management skills
  - Skills in bridging data science and marketing
  - Knowledge on legal aspects of data
- Organisational capabilities:
  - Organisational culture for open data OI
  - Ability to hire, retain and integrate a mix of talent
  - Ability to engage with OD publishers and consumers
  - Capability for continuous innovation

OD advantages for SMEs
- Integrating OD with closed data
- Access to new, previously unavailable data
- Proprietary data replacement
- Scaling-up through coupled OI

Inbound OI
Coupled OI
### Appendix 1: Research Participants

<table>
<thead>
<tr>
<th>Case</th>
<th>Organisation Type</th>
<th>Sector(^1)</th>
<th>Interviewee</th>
<th>Publish OD (outbound OI)</th>
<th>Consume OD (inbound OI)</th>
<th>Face to Face/Skype/Telephone interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Corporation</td>
<td>Engineering</td>
<td>Product/innovation manager</td>
<td>No</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>2</td>
<td>SME</td>
<td>Retail</td>
<td>Product/innovation manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>3</td>
<td>Public agency</td>
<td>N/A</td>
<td>Head of Data</td>
<td>Yes</td>
<td>No</td>
<td>Face to face</td>
</tr>
<tr>
<td>4</td>
<td>SME</td>
<td>Business services</td>
<td>Owner-manager</td>
<td>No</td>
<td>Yes</td>
<td>Telephone</td>
</tr>
<tr>
<td>5</td>
<td>Foundation</td>
<td>Business services</td>
<td>Head of Data</td>
<td>No</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>6</td>
<td>Public agency</td>
<td>N/A</td>
<td>Head of Data</td>
<td>Yes</td>
<td>No</td>
<td>Telephone</td>
</tr>
<tr>
<td>7</td>
<td>SME</td>
<td>Business services</td>
<td>Director</td>
<td>Yes</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>8</td>
<td>SME</td>
<td>Business services</td>
<td>Director</td>
<td>Yes</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>9</td>
<td>SME</td>
<td>Business services</td>
<td>Director</td>
<td>Yes</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>10</td>
<td>Public agency</td>
<td>N/A</td>
<td>Product/innovation manager</td>
<td>Yes</td>
<td>No</td>
<td>Face to face</td>
</tr>
<tr>
<td>11</td>
<td>SME</td>
<td>Consultancy</td>
<td>Owner-manager</td>
<td>No</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>12</td>
<td>SME</td>
<td>Consultancy</td>
<td>Owner-manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>13</td>
<td>Public agency</td>
<td>N/A</td>
<td>Head of Data</td>
<td>Yes</td>
<td>No</td>
<td>Skype</td>
</tr>
<tr>
<td>14</td>
<td>SME</td>
<td>Software development and services</td>
<td>Owner-manager</td>
<td>No</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>15</td>
<td>SME</td>
<td>Retail</td>
<td>Product/innovation manager</td>
<td>No</td>
<td>Yes</td>
<td>Telephone</td>
</tr>
<tr>
<td>16</td>
<td>SME</td>
<td>Real estate</td>
<td>Owner-manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>17</td>
<td>SME</td>
<td>Leisure</td>
<td>Owner-manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>18</td>
<td>Public agency</td>
<td>N/A</td>
<td>Product/innovation manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>19</td>
<td>SME</td>
<td>Software development</td>
<td>Director</td>
<td>Yes</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>20</td>
<td>SME</td>
<td>Business services</td>
<td>Product/innovation manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Telephone</td>
</tr>
<tr>
<td>21</td>
<td>SME</td>
<td>Software development</td>
<td>Software developer</td>
<td>No</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>22</td>
<td>SME</td>
<td>Business Services</td>
<td>Director</td>
<td>Yes</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>23</td>
<td>SME</td>
<td>Consultancy</td>
<td>Owner-manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>24</td>
<td>SME</td>
<td>Software development</td>
<td>Owner-manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>25</td>
<td>SME</td>
<td>Business Services</td>
<td>Director</td>
<td>Yes</td>
<td>Yes</td>
<td>Telephone</td>
</tr>
<tr>
<td>26</td>
<td>Corporation</td>
<td>Retail</td>
<td>Product/innovation manager</td>
<td>No</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>27</td>
<td>SME</td>
<td>Consumer</td>
<td>Product/innovation manager</td>
<td>No</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
<tr>
<td>28</td>
<td>Corporation</td>
<td>Consultancy</td>
<td>Director</td>
<td>No</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
</tbody>
</table>

\(^1\) Sector information is not provided for government departments as this would identify the organisations and potentially the interviewees, violating participant anonymity.
<table>
<thead>
<tr>
<th>No</th>
<th>SME</th>
<th>Consultancy</th>
<th>Position</th>
<th>Yes</th>
<th>Yes</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>SME</td>
<td>Consultancy</td>
<td>Director</td>
<td>Yes</td>
<td>Yes</td>
<td>Skype</td>
</tr>
<tr>
<td>30</td>
<td>SME</td>
<td>Consultancy</td>
<td>Owner-manager</td>
<td>Yes</td>
<td>Yes</td>
<td>Face to face</td>
</tr>
</tbody>
</table>