

Analytical Report n10



Analytical Report 10: Open Data and Entrepreneurship



Analytical Report 10: Open Data and Entrepreneurship

Last update: February 2018

www: <http://www.europeandataportal.eu/>

@: info@europeandataportal.eu

Licence: CC-BY



Authors: Johanna Walker and Elena Simperl (University of Southampton)

Reviewers: Wendy Carrara and Cosmina Radu (Capgemini)

Contents

1. Introduction	4
2. The Role of Entrepreneurs in Open Data Innovation	5
3. Open Data Barriers to Entrepreneurship.....	7
4. Entrepreneurship and Open Data Sustainability	9
5. EU Open Data Policy and Entrepreneurs	11
6. The Impact of Entrepreneurs and Open Data.....	13
7. Best Practice in Open Data Start-Up Support – An Analysis.....	15
8. Recommendations	20
9. Conclusion.....	22

1. Introduction

The mechanisms of value of Open Data operate on multiple levels. Multiple Open Data benefits can be outlined, including transparency, increased government effectiveness and efficiency, increased participation and engagement and the release of social and commercial value through innovation.¹

This economic value is created through direct and indirect benefits. The European Data Portal Extensive Study, 'Creating Value through Open Data' identifies direct benefits as revenues, Gross Added Value, job creation and cost savings². Among the indirect benefits are innovations (new goods and services), time savings, the growth of the knowledge economy, increased efficiencies of public services and the growth of related markets.

The scale of this is substantial. The European Data Portal Analytical Report, 'Economic Benefits of Open Data' estimates the cumulative market size for the period 2016-2020 at between 1.138 billion euros and 1.229 billion euros. By 2020 there will be slightly less than 100,000 Open Data jobs in Europe, creating almost 20,000 new direct Open Data jobs by 2020 as compared to 2017.³

Aside from the financial value, Open Data products and services may benefit our congested cities by improving the way we interact with public transport, or delivering better public transport use, or enabling more sharing of private transport. They may ensure food security by matching soils to crops, or producers to buyers. They may increase our health by enabling us to find the right fitness classes or speed up the process of diagnosis or treatment when we are not well. They frequently ensure that the information we need is at our fingertips when we need it in a format we can use. There are many, many potential advantages.

The benefits of geospatial products alone are extensive. Sentinel open earth observation data is estimated to provide 10.8 – 13.5 billion euros in cumulative value across Europe by 2020.⁴ The socio-economic value of open geodata in Denmark alone was estimated at 3.5 billion Danish krona (470 MEUR) in 2016, representing a doubling in efficiency and production gains since 2012. In November 2017 the UK announced a new Geospatial Commission, which will launch a drive the move to use the data more productively, unlocking potentially up to £11bn (12.5 billion EUR) of extra value for the economy every year. Open Data can have a positive impact in many other business-friendly ways. It cuts costs, removes the need to manage and protect copyrights and reduces the effort required to foster relationships with data publishers.

That there is potential for entrepreneurial development of innovative, economically beneficial products and services from Open Data makes logical sense, or as a 2016 NESTA report phrases it, it is 'intuitively plausible'⁵ as data is increasingly available for the creation of new insights and activities. Although innovation with open data takes place across all sizes and ages of organisations, entrepreneurs and start-up businesses are important players. A report on the start-up companies in the Open Data Incubator for Europe (ODINE) found that they "contribute to the development of an

¹ <http://opendatahandbook.org/guide/en/why-open-data/>

² Carrara, W, Chan, W S, Fischer, S and van Steenberg, E (2015) European Data Portal Extensive Study, '[Creating Value through Open Data](#)'.

³ Berends, J, Carrara W and Radu, C (2017) European Data Portal Report. [Analytical Report 9: Economic Benefits of Open Data](#).

⁴ [Socio-Economic Impact of Copernicus in the EU 2016](#)

⁵ Rubenstein M, Cowls, M and Cath, C (2016) [Open Data Innovation Working Paper: An International Journey to Discover Open Data Innovation](#).

Open Data ecosystem in Europe covering all segments of the data value chain,” particularly in more innovative segments.⁶

There is no commonly agreed definition of what constitutes a start-up. However, youth, size and the potential for growth all contribute to the concept of a start-up. A valid working definition could be “a company that is less than 5 years old, has fewer than 100 employees and is highly scalable”. This last aspect is critical; a café in town may meet age, employee (and turnover, if used) criteria, but may never intend to grow to more than one site. A start-up, on the other hand, is looking to grow, fast and efficiently.

Open Data 200 Italy can help us understand the relationship between Open Data and companies of different ages, and how Open Data is particularly valuable as a raw material for start-ups. It is a joint project designed and implemented by the GovLab in partnership with the Fondazione Bruno Kessler, an international research institute located in Trento. It conducted the first comprehensive, internationally comparable study of Italian companies that are using Open Data to generate business, develop products and services, and create social value. Of the 55 companies listed so far, 19 have been created in the last 3 years. Twenty-seven - nearly half - have been founded since the launch of the Italian Open Data portal, dati.gov.it, in late 2011.

Established companies (Over 10 years old) working with Open Data:	19
Companies 5 - 10 years old working with Open Data:	12
Companies under 5 years old working with Open Data:	27

2. The Role of Entrepreneurs in Open Data Innovation

Open Data appears to be a valuable raw material but its benefits are realised mostly through transformation, analysis, aggregation and synthesis. A quantitative analysis of measurements from 61 countries found that openness “*positively affects the ability of society to generate value from data through the innovation mechanisms*” and that “*data-driven innovation positively affects value through generation of new knowledge, new processes, services and products, and new businesses*”.⁷ The second of these hypotheses was particularly statistically significant (Jetzek et al, 2014). The creation of new products and services is just one of many potential methods of adding economic value. There is also the possibility of improving business processes and intelligence (especially in targeted marketing and benchmarking areas), creating opportunities to better match supply and demand and enabling collaborative opportunities.

Data publishers are not ideally set up to discover new value in their own data. This may be because of cost, competing calls on time, a lack of understanding where that value might lie, or simply because the processes and skills required for the production and collection of data are not those required to exploit the data. While a substantial share of Open Data is published by national, local or municipal government, Chris Yiu, in a 2012 paper entitled, ‘Where Next for Open Data?’, has argued that the exploitation of Open Data is not an appropriate function of government, as this deters competition

⁶ IDC (2016) [Impact Assessment of the ODINE Programme](#)

⁷ Jetzek, T, Avital, M and Bjorn-Andersson, N (2014) [Data-driven Innovation Through Open Government Data](#).

and weakens innovation⁸, and Maurice McNaughton of the Caribbean Open Institute states that engaging the private sector in re-usability and innovation is critical.⁹ Early forays into app creation by Transport for London (TfL), the public transport authority in London in the United Kingdom, were soon superseded by private companies such as the early-stage, venture-backed Citymapper. In 2015 over 360 apps used TfL data, growing to over 600 in 2017. The wider implication for Transport for London, therefore, is that they should not invest in training and staff to develop apps; instead, they can focus more resources on the activities they are best placed to undertake: publishing data well. In this way, they can have a greater effect via the use of the data by other organisations. In fact, TfL calculate they have a 58:1 return on their investment in Open Data.¹⁰ On a much smaller scale, Code4Vilnius engages citizens of all kinds in using the Lithuanian city's Open Data to address issues such as increasing recycling rates and carpooling. Such efforts are integrated with the municipality, but crucially, not executed by it.

The UK government records the category of the person or organisation making a request for datasets on data.gov.uk. After private individuals, start-ups and SMEs¹¹ are the largest categories of requesters. One hundred and nineteen start-ups have made requests for 55 datasets, and 191 SMEs have made requests for 89 datasets since 2014. A further reason to believe that Open Data and early-stage, entrepreneurial companies are particularly well-matched, is that they have the required flexibility to combine the social responsibility aspects of Open Data, such as transparency and participation, with the creation of profit. This gives them a distinct motivational difference from developers or civic technology activists who may not aim to exploit the product for direct financial benefit. This profit motive is important for a number of reasons, including the sustainability and economic growth of the start-up.

However, it is not just entrepreneurs that are important to open data, but open data is also important to entrepreneurs. Open Data has an enabling effect – it is a driver for entrepreneurship, creating opportunities for new business models, products and services, by providing a wealth of free material for building apps, websites and software. For instance, the Paris DataCity initiative set challenges that could be addressed with over 40 open and private datasets from both the city and corporate partners. This attracted over 400 start-ups.¹² The impact assessment of the Open Data in Europe (ODINE) accelerator found that “advanced Open Data markets provide a favourable environment for innovators with new business ideas.”¹³

Similarly, the European Data Portal Analytical Report “Digital Transformation and Open Data” noted that almost half the re-users identified in their research were considered “enrichers” - organisations using Open Data to gain new or better insights they can use to deliver services or products to their customers - often completely new services which could not exist without Open Data.

⁸ Yiu, C (2012) [Where Next for Open Data](#).

⁹ McNaughton, M (2014) [Harnessing Open Data to Achieve Development Results in Latin America and the Caribbean](#)

¹⁰ <http://www.gov.scot/Publications/2013/12/6550/4>

¹¹ These are self-determined categories.

¹² <https://datacity.numa.co/paris/en/program/>

¹³ IDC (2016) [Impact Assessment of the ODINE Programme](#)

Case Study: Synergise (Slovenia) ¹⁴

Some of the most valuable and voluminous datasets in Europe, with almost limitless applications, are also some of the most complex. This means that, despite several terabytes of free and open earth observation data being generated every day, much of it is hugely underutilised. Synergise's product, Sentinel Hub, provides access to quadrillions of bits of information coming from Copernicus, USGS and other satellites in a matter of seconds, making them easily accessible to end-users and application developers around the world. In this way, Synergise plays an important intermediary role, both supporting the data owners to publish data in a more accessible way – just like TfL did in the previous example - but also by enabling other start-ups to focus on creating added-value products and services instead of data cleaning and management. By doing this, Synergise increases the uptake of data from the Copernicus constellation of satellites, which make millions of daily observations, as well as from a global network of thousands of land-, air- and marine-based sensors.

Synergise originated in Slovenia, however, the company has clients both within and beyond Europe including Croatia, Serbia, Macedonia, Montenegro, France, the United Kingdom, Azerbaijan, Moldova, Ghana, Nigeria, Tanzania and Mauritius. They make their data available to all customers under certain conditions and charge businesses, governments, academia and professional users a subscription fee for premium elements.

3. Open Data Barriers to Entrepreneurship

The entrepreneur as a special person who has a greater propensity to risk than average is a kind of business mythology, although the evidence seems evenly ranged for and against this belief. However, uncertainty and potential loss are an inevitable part of early stage business. Open Data can facilitate the reduction of risk, by removing the time and cost friction of locating, negotiating and paying for data, but there remain substantial barriers for re-use. The European Data Portal's Analytical Report 'Barriers in Working with Open Data' finds that these are present across organisational, technical, financial, political and legal domains.¹⁵ Particularly of concern were the specific barriers associated with geospatial data: as mentioned above this is possibly the most valuable and re-used type of Open Data, however it requires particular knowledge and skill to use.

As Open Data is readily available at minimal cost, entrepreneurs are able to use it both as an asset in its own right, and, as explored in the next case study, as a placeholder that can help establish the business until it is able to create, or pay for, a more suitable dataset. However, there are specific challenges of using Open Data that are particularly pertinent to entrepreneurs. If a dataset is not already open, then it can be a lengthy process to build a business case and request the data. Unless the entrepreneur has the means to support the business during this process, there is a real risk of running out of funds before the data becomes available. Hackers, democracy advocates and corporates, among others, may be equally reliant on Open Data to create their product or service, but perhaps not so reliant on its producing a revenue stream.

Other concerns include financial agreements between large corporations and government departments creating a disincentive for the opening of valuable datasets; a lack of standardisation

¹⁴ The Case Studies are based on IDC (2016) Impact Assessment of the ODINE Programme

¹⁵ Berends, J, Carrara W an Vollers, H (2017) European Data Portal [Analytical Report 5: Barriers in Working with Open Data](#).

and gaps between a government's pro-Open Data stances and actual availability of datasets. Entrepreneurs may also suffer from not being as well-networked with data publishers as larger competitors and conversely, publishers may not be sufficiently engaging with open data start-ups to ensure data publication and prioritisation in line with their needs. Both the public sector and commercial organisations can also lack an open data reuse culture. This might take the shape of a belief that sharing the data will automatically mean a transfer of the value away from the data owner, or simply little understanding of how value is created from open data. The European project Data Pitch is engaging with exactly this issue and aims to show how innovation and economic value can be created for all by sharing data.¹⁶

Accessible, easy-to-use data portals have a vital role to play here, in facilitating the discovery and use of (primarily government) data and promoting a wide variety of societal objectives including economic growth and the creation of new innovative products, services and apps. More detail can be discovered in the European Data Portal's Analytical Report, 'The Future of Open Data Portals'.¹⁷

As well as portals, finance and networking, a key issue is the chilling effect of 'share-alike' licensing. Share-alike licensing has the function of promoting crowdsourcing and stimulate collective data generation and sharing. This can be illustrated with a well-known example, that of Open Street Map (OSM). OSM has always had a 'share alike' restriction, which states, "You can basically do whatever you want with our data inside your own home or organisation, but if you then publish your results as data, then you need to tell folks about us (attribution), and share back to the public any improvements that you made, (share alike)."

This is slightly more restrictive than 'freely used for any purpose'. OSM spell out the implications on their Legal FAQs¹⁸, which state, "you can charge any amount of money you want for any service or data you provide. However, since the data that is derived from OSM data must be licensed as above, other people may then redistribute this without payment." This can make entrepreneurs avoid such datasets. A start-up company with limited resources to spend on exploring licensing issues may simply decide to avoid those datasets, and this may result in lost opportunities for entrepreneurs. The size of the potential impact of lost opportunities from licensing, financial, portal or other issues is illustrated by a 2017 survey of start-ups associated with the Open Data Incubator for Europe (ODINE).¹⁹ Sixty-three percent of respondents said Open Data was a key ingredient in their products or services, with almost a quarter stating that it formed the core of their products and services (24%). Only 13% believed Open Data to be marginal, non-essential ingredient in their offering. This suggests very clearly that without Open Data, these businesses would not exist.

Case Study: PlumeLabs (Fr)

Plume Labs is an example of an entrepreneurial company that was able to achieve minimum viable product and prototype with Open Data. Flow, Plume Labs' smart, connected mobile accessory, helps users track and monitor air pollution wherever they are, in order to help them understand their environment and monitor exposure to air pollution for improved wellbeing. Their free mobile app, the Plume Air Report, allows users to find clean air by providing real-

¹⁶ datapitch.eu

¹⁷ Simperl E and Walker, J (2017) European Data Portal [Analytical Report 8 The Future of Open Data Portals](#)

¹⁸ http://wiki.openstreetmap.org/wiki/Legal_FAQ

¹⁹ IDC (2016) [Impact Assessment of the ODINE Programme](#)

time air quality forecasts in over 60 countries thanks to Open Data and machine learning predictive algorithms.

Plume Labs' long-term vision is to crowdsource air quality levels but chose to start building its global air quality platform on top of Open Data from public sector environmental monitoring networks. Although the sensors can produce more personal exposure data, and be more sensitive to urban air pollution variations, getting to scale with the sensors is a lengthy and expensive process, and Open Data has enabled Plume Labs to come to market with a product without the necessity for a wide installed base of Flow users. Plume plans to release more Open Data for researchers and activists leading the fight against air pollution, which also reflects their commitment to social responsibility.

4. Entrepreneurship and Open Data Sustainability

Entrepreneurship is key for the sustainability of Open Data, in two senses. The first aspect of sustainability is a continuing stream of funding for the publication of Open Data, whether by government or the private sector. While economists such as Open Knowledge founder Rufus Pollock have stated the macro argument for the benefit of Open Data, the micro-economic arguments remain confounding.²⁰ If Open Data is both free as in speech and, due to the almost zero cost of replication over the web, free as in beer, how is revenue acquired? And without revenue, who is responsible for the costs – which are, in some cases, substantial? Governments at the national, regional and municipal level have largely funded this initial step, but these pockets are not endlessly deep. The European Commission in its Handbook on Sustainability of 2006 states, “a project is sustainable when it continues to deliver benefits to the project beneficiaries and/or other stakeholders for an extended period after the EU’s financial support has been terminated.”²¹ In other words, the way that the vast majority of Open Data is currently made available, through portals, is not yet sustainable, and entrepreneurship has a role to play in this.

The first is as a direct creator of macro value – the transformational element that allows Open Data to positively impact on taxation and employment by creating viable businesses with job opportunities. This requires that the Open Data business itself is sustainable in terms of revenue and audience. The Europe-wide accelerator Data Pitch asks start-ups and SMEs to solve a challenge with data, but also requires that solution to be a product or service with a wider market, in order to ensure sustainability of the solution.

The second aspect of sustainability is that which is associated with the creation of value across not only financial but also social and environmental dimensions – the triple bottom line. In the Open Data Maturity in Europe Report the social impact of Open Data on society is measured on two different levels: to what extent Open Data has a positive impact on the

²⁰ Pollock, R (2008) [The Economics of Public Sector Information](#).

²¹ European Commission (2006) [Sustainability of international cooperation projects in the field of higher education and vocational training](#).

environmental sustainability in a country and to what extent Open Data has a positive impact on the inclusion of marginalised groups in a country.¹

We have already mentioned how this social, sustainable aspect seems ingrained in the DNA of many of the Open Data start-ups which are not purely focused on data analysis or handling. These businesses have twin goals: to earn revenue and to achieve social, cultural, community economic and/or environmental outcomes. This suggests an inherent synergy between social enterprise and Open Data. This may be, in part, because the socially beneficial nature reduces potential for criticism over the use of public sector information for commercial use. One of ODINE's key objectives was to select and accelerate enterprises able to achieve a "triple bottom line" including economic, social and environmental benefits.

Open Data start-ups are also playing a role within the start-up ecosystem to promote Open Data and thus contribute to its ongoing sustainability. Northern Ireland company SeeSense produces light sensors for cyclists, which connect via Bluetooth to an app on Android phones. The app transmits anonymised data on the cyclist's environment - such as the quality of the road surface, light levels, as well as cycling routes, collisions and near-miss events to telecoms company BT's CityVerve hub in Manchester and also with Dublin City Council which is then made available for innovators and planners.

Case Study: Farmdog.ag

Farm Dog is making healthy and environmentally sustainable food accessible to everyone, changing the way growers choose, use, and purchase pest and disease treatments. Treatment efficacy reviews, regional alerts, and variable rate treatment application provide growers with a new toolkit to manage their fields.

Farm Dog was founded in 2015 and now has offices in Israel and California. The company takes the most advanced tools developed for the largest of farms and tailor-fits them for small and medium-sized farms. The Farm Dog platform is built on a foundational layer of in-field sensors combined with multiple data streams such as farmer inputs, micro-local weather, aerial imagery, and agronomist alerts. The farmer then receives real-time monitoring and alerts of his field to help him or her optimise resource use, increase yield, and better utilise managerial time.

The founders explicitly wanted to create a new venture that was both challenging and did "good" (in other words, not marginal improvements on e-marketing campaigns). Farmdog has been funded by the European Commission's ODINE Programme, Microsoft Ventures and an Israeli VC, all of whom saw the commercial as well as societal potential.

5. EU Open Data Policy and Entrepreneurs

As of 2017, all EU countries had Open Data policies which encouraged re-use, although these vary in their development.²² Governments that open up data may gain financially through two channels: increased employment leading to lower unemployment subsidies and higher tax revenues, and higher indirect tax revenue from related products and services. Additionally, the public sector benefits from significant efficiency gains and reduced transaction costs. The European Data Portal's Analytical Report 'Economic Benefits of Open Data' estimates that 25,000 jobs will be created by Open Data in 2020, and more than 30 million euros of public administration savings will be made in 11 countries.

Further, in the words of the W3.org paper, 'Best Practice: Establish an Open Data Ecosystem': "Citizens are not interested in data: they are interested in services being built with the available data and information."²³ Promoting the use of Open Data and providing explicit support for entrepreneurs is therefore an important role of government. This support might take several forms: reduction of regulatory constrictions, increasing networks or funding via innovation and technology grant programmes. Among these, there are some key specific policy actions that can be explored.

Competitions to facilitate innovation are a cornerstone of Open Data support policy. These often take the form of 'hackathons', inexpensive and time-bounded opportunities for developers to work with data to develop innovative products and services, often along a theme, such as 'Mobility' or 'Energy'. Such promotion of re-use may be run by governments, such as Stockholm Open Lab, start-up groups, such as Belgian co-working space Betacowork's Open Data Hackathon, by a combination of academia, activists and government, such as Apps for Gent, or by corporates, such as German transport operator Deutsche Bahn, which runs regular hackathons with its Open Data in Berlin. These events can be the catalyst for people and ideas to come together to create a firm. Energy saving big data processing firm Mastodon C was created after an initial Open Data hackathon weekend.²⁴

However, there have been serious challenges to the hackathon as a method for supporting the development of Open Data start-ups. Studies in 2011 and 2013 discovered that, despite some interesting products being created during the event itself, very few sustainable businesses emerged from hackathons.²⁵

There are various possible reasons for this. One aspect is the motivation for participation in hackathons. Some people become involved to learn more about coding and data rather than to create sustainable innovation, as discussed above. There can be a lack of ongoing engagement or resources to sustain the ideas. However, entrepreneurs are well-placed to organize knowledge and resources to take advantage of potentially valuable opportunities, and as such are a key part of the Open Data ecosystem that needs to be in place to ensure hackathons are productive in the long term. Further, even if initial prototypes are abandoned as unfeasible, entrepreneurs often pragmatically move to 'plan B', which may or may not be based on the same data but will certainly include some of the learning the entrepreneur has accumulated from the process.

²² Berends, J, Carrara W and Radu C (2017) European Data Portal. [Analytical Report 9: Economic benefits of Open Data](#)

²³ W3 Working Group (2016) Best Practice: [Establish an Open Data Ecosystem](#).

²⁴ <https://theodi.org/case-studies/mastodon-c-case-study>

²⁵ Martin, S, Turki, S, Foulonneau, M and Ijahadene, M (2013), [Risk Analysis to Overcome Barriers to Open Data](#).

A more extensive version of the hackathon is the ‘challenge’, such as Open Data FWD in Amsterdam (2014) the Copernicus Big Data Challenge (2017), the Open Energy Data Challenge in Vienna (2017) and the Open Data Challenge Series in the UK (2013 – 15). In this last Challenge, an original funding of £1.2m was used to support teams and individuals using Open Data to generate new and sustainable solutions to key social challenges. It is predicted that, over the three years since the programme ended, it will inspire a return of £5 - £10 to the UK economy for every £1 invested (£5.3m to £10.8m of Gross Value Added). The impact report also estimated that between 75 and 141 jobs would be created alongside the potential to generate further social and environmental impacts.

Another approach, which is widely used in supporting start-ups across the globe, is the incubator or accelerator. There have been a number of these across Europe, including FINODEX, ODINE, Data Pitch (which combines the challenge format with an accelerator) and the OpenActive accelerator (May 2018). The EU is investing strongly in this area, with 7.1m euros of funding for the Data Pitch accelerator, which encourages corporates and start-ups to collaborate on data, as a follow on from ODINE. Eighteen data start-ups entered the Data Pitch accelerator in February 2018.

The Open Data Incubator for Europe (ODINE)²⁶ attracted and funded a group of innovative digital start-ups and very young companies, plus a few SMEs looking for opportunities for growth. For most of the companies the project launched with ODINE represented their core business, suggesting that the programme’s influence on their business and growth perspective was strong. An impact report on ODINE concluded it had succeeded in inspiring and promoting a range of new business ideas highlighting the value added of Open Data in the data market. Each of the 57 start-ups and SMEs leveraged two or more types of Open Data, with a strong concentration of interest in geospatial/ mapping and environmental data. This suggests that proactive policies improving the usability and availability of open data sets are likely to stimulate private initiatives for the exploitation of data in a positive virtuous cycle²⁷. The number of start-ups applying to ODINE (over 950) and fellow Open Data accelerator FINODEX (493) highlights the size of interest from entrepreneurs in applying to such programmes.

No significant correlation emerged between specific offerings and growth dynamics, demonstrating that there is not a single way to success for these companies. However, a positive correlation between national maturity of the Open Data market and the number of ODINE successful applicants by country points out that a rich Open Data environment provides favourable conditions for innovators in this field²⁸. This means that proactive policies improving the usability and availability of Open Data sets are likely to stimulate private initiatives for the exploitation of data in a positive virtuous cycle.

Companies based on inbound Open Data are also now publishing their own data as open, thus creating a ‘second generation’ of Open Data. Start-ups doing this understand that value is created by what is done with Open Data, rather than existing inherently in the data itself. Sports technology company Imin, a participant in the ODINE accelerator, subsequently joined the data sharing accelerator Data Pitch as a data provider with the aim of supporting other companies to work with Open Data and to promote the paid-for data services they develop. Further, Imin became the lead organisation

²⁶ opendataincubator.eu

²⁷ IDC (2016) [Impact Assessment of the ODINE Programme](#)

²⁸ Ibid.

of Openactive.io, a community of sector organisations advocating Open Data and collaborating on open standards.

CASE STUDY: Xpressomics (Estonia)

Xpressomics is a search engine for genetic data. When biologists perform experiments they sequence many organisms – writing down the genetic code for those organisms. That code is represented as data and the amount of this data is growing fast. The problem is that the data is scattered all across various public databases and it is difficult to search for relevant and important information about genes from this data. Xpressomics takes this publicly available genetic data and indexes it. Much like Google, they rank the data according to relevancy. In this way, they enable science to extract full value from the data.

Xpressomics was supported by the Finodex accelerator, which specifically focused on supporting companies using Open Data with funding, training and access to the Fiware technology platform. They credit the accelerator with making their development process more efficient, as well as making it easy to procure additional technical resources and reach a minimum viable product.

6. The Impact of Entrepreneurs and Open Data

Although case studies such as those cited in this report abound, there is as yet a lack of systematic measurement via impact metrics, and our understanding of how innovation occurs through Open Data activities is at an emergent stage. Analysis of Open Data use by entrepreneurs is challenging, as there is no set framework or methodology for evaluating use.

There is no requirement for the reporting on the use of Open Data, nor an agreed way of identifying or tracking it. Common Assessment Methods is one project that aims to tackle this and other related issues²⁹. Even were this to be facilitated, counting the number of businesses built around open data is not a sufficient metric to gauge impact as the vast majority of the economic benefit will accrue in small gains spread across a wide range of firms.

Tracking is vital because, in the long term, it is only by evidencing beneficial use that organisations (even public sector ones) can justify the ongoing costs of publishing data. One of the ways to track how data is being used is via Application Programming Interfaces (APIs) with keys. Once an organisation – particularly a corporate one – is committed to validating and issuing APIs and then tracking use, they may find that this requires more capacity and resource than they can actually afford, so they need to reduce the number of companies (including start-ups) using their API. Thus, what is genuinely meant to be Open Data starts operating rather like “named access data” as this is the most efficient way for some organisations to publish. Therefore, it is important to keep a balance between the commercial imperatives and the more altruistic perspectives.

²⁹ <https://webfoundation.org/2014/06/towards-common-methods-for-assessing-open-data/>

However, two good, quantitative sources of evidence of impact are the 57 start-up companies in the Open Data Incubator for Europe (ODINE) and the 40 start-up companies in the FIWARE Open Data accelerator (FINODEX). Impact studies from ODINE found that combining the Open Data businesses with the acceleration resulted in an estimated 110 €M of cumulative revenues for the 57 supported companies in the years 2016-2020 with a final projection of 784 jobs created over this period. In FINODEX, 50 new jobs were created by their top 20 companies. The companies overall delivered a return on investment of 75%.

Researchers studying the ODINE companies found they use on average two or more types of Open Data – in other words, they are highly engaged with it. This is because they are focusing on multiple markets; because they want to improve their competitiveness with a strong value proposition, and lastly, because, like the above-mentioned example of Imin, they are conscious of the potential economic and social impact of Open Data on the ecosystem, and they want to contribute to it.

The younger companies amongst the ODINE cohorts more often combine different typologies of Open Data for their solutions, while mature companies appear to be more focused on specific typologies of Open Data. Extrapolating from this, it might be expected that as new start-ups enter the market they will increasingly break down the Open Data “silos” and accommodate multiple data inputs. Companies were analysed in terms of how many clusters of Open Data they use and which of them. Researchers found a high concentration of companies using the Environment Open Data cluster including weather, agriculture and geospatial data (approximately half of them), followed by the Vertical Markets cluster (including housing, manufacturing, transportation and finance, among others) and the Social cluster (including education, demographics and scientific research. In the same context, the European Data Portal Study “Re-using Open Data,” identifies the most and least used data set categories across all types of business, based on a survey methodology³⁰. The results are very similar to those found in the ODINE survey. The EDP plans to develop more automated tracking methodologies in the future.

Case Study: OpenCorporates (UK)

OpenCorporates was founded in 2010, and boasts the largest open database of companies in the world. It has enjoyed a high growth environment with increasing revenue, employees, investment, and customers, including a 25% year on year compound growth on own funds. It has recently launched a new product, OpenGazettes. Government gazettes are used in many parts of the world to publish important notices about companies, from incorporation and dissolution to winding up orders, AGMs, mergers and more. However, these notices are little known, and rarely connected with the companies to which they relate.

OpenGazettes/OpenCorporates is currently doubling its offering; working with the EU Horizon 2020 project euBusinessGraph, developing a new ‘corporate events’ functionality, targeting at least 40% year on year revenue growth. Customers include traders, businesses, and insurance companies, law enforcement, investigators and reporters. Whether assessing Open Corporates by the innovativeness of its services, financial growth or connectedness up and down the value chain, it is making an impact.

³⁰ Carrara W, Vollers H: European Data Portal. [Reusing Open Data](#).

7. Best Practice in Open Data Start-Up Support – An Analysis

We examined the corpus of documents collected on the European Data Portal (EDP) to understand more about the focus of Open Data entrepreneurs along the dimensions discussed here and any others. The EDP aims to act as a “One-Stop-Shop” for Open Data information at European level and provide a repository for all relevant EU Open Data articles and documents, and is therefore an appropriate location to discover best practice in this domain. In total, there are 714 documents that have been collated or generated by the EDP.

We initially searched for the following keywords and found that the term ‘start-ups’ was the most utilised by reports, documents, blog posts and other articles on this topic collected by the EDP (Figure 1).

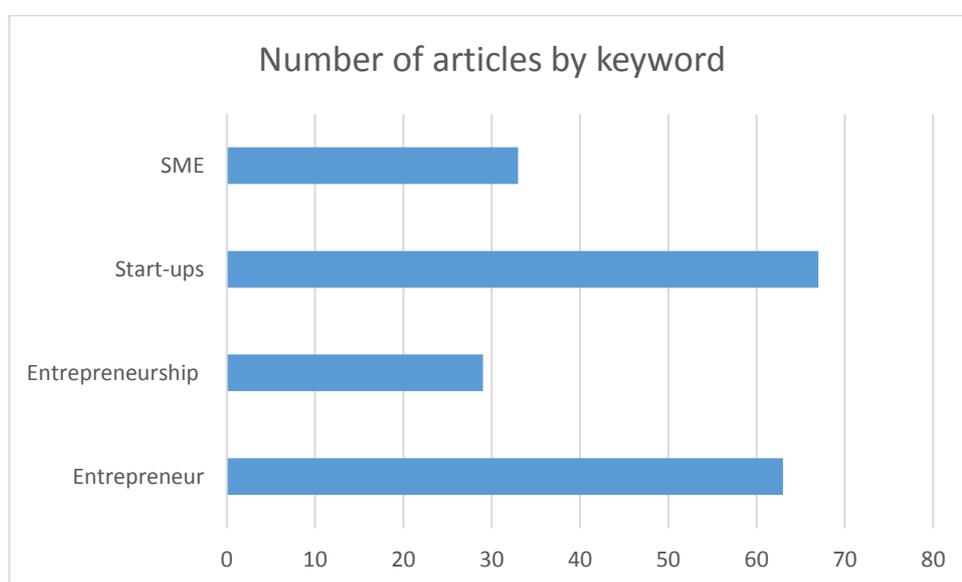


Figure 1 Documents on the EDP containing key terms

Therefore, we chose this sub-group of documents for our initial analysis. Demographically they reported on activities from 14 European countries, but over half the documents were pertinent on a Europe-wide basis (Figure 2).

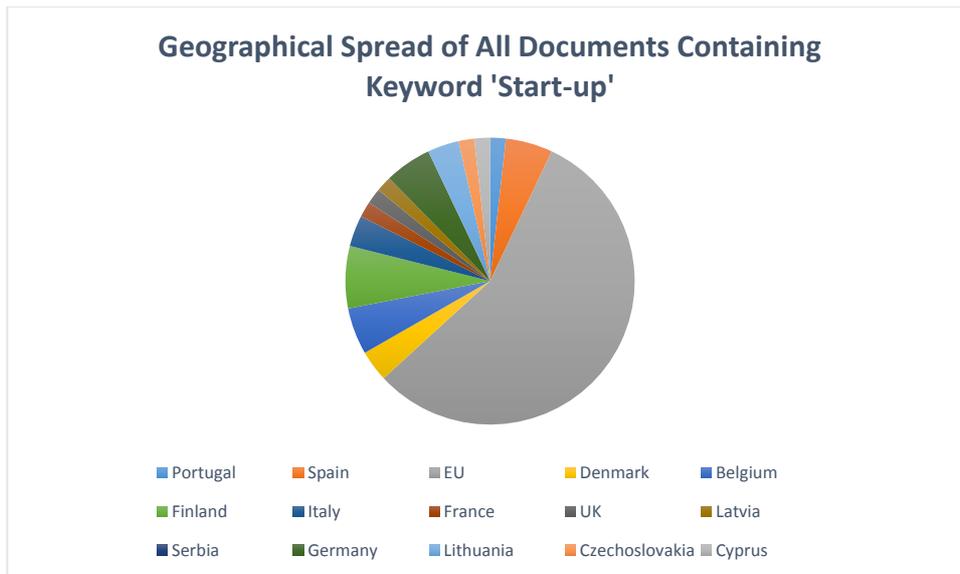


Figure 2 Geographical spread of all documents containing the keyword 'start-up'

In terms of the content, over half of the documents referred to start-ups as one amongst many stakeholders, including citizens, journalists, public sector, third sector, large businesses and more. We therefore focused on the documents which made start-ups their key focus. These fell into three groups: articles about competitions; re-use examples and the largest group, with 13 documents focused on support for start-ups (Figure 3).

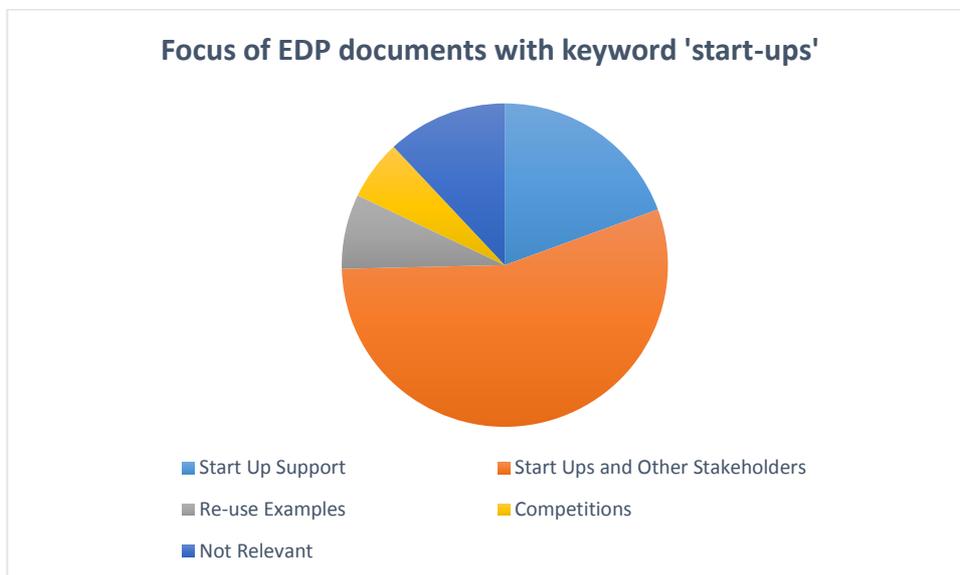


Figure 3 Thematic focus of documents containing the keyword 'start-up'³¹

³¹ European Data Portal (2014). Best Practice: [Setting Up an Open Data Business.](#)

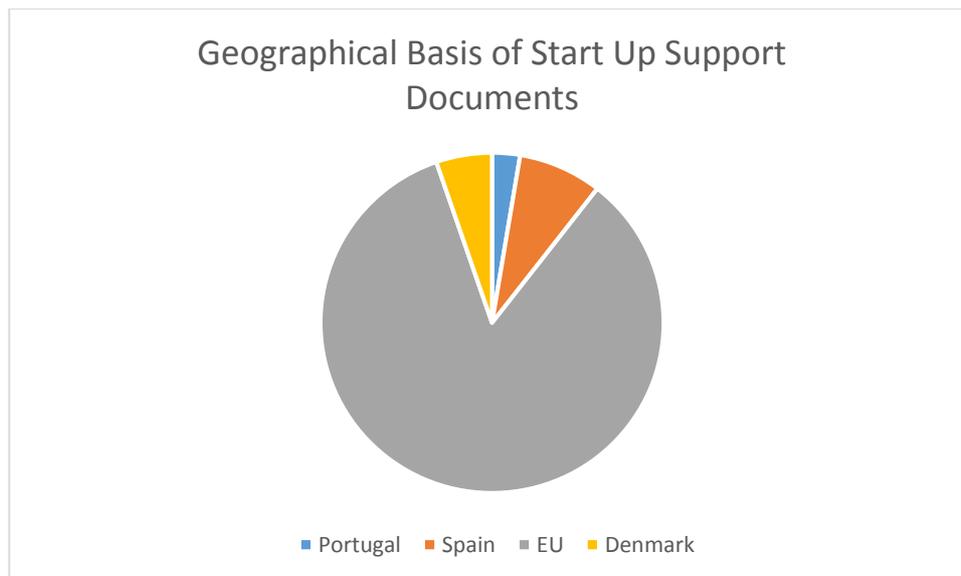


Figure 4 Geographical spread of documents focused on start up support

The majority of the information under analysis is based on EU wide learning and initiatives (Figure 4). The support described in the articles covers a number of different activities and approaches, from experiential descriptions to surveys to outlines of approaches. There were a couple of very specific niches but almost all the articles referred, either explicitly or implicitly, to the vital role a strong Open Data ecosystem plays in supporting start-ups. A second strong theme was sustainability and the role that commercial start-ups had in creating this. In practice, these two themes are combined, as a strong ecosystem offers the best opportunities for open data re-use, while proven and beneficial re-use is required to support the publication of open data. In the rest of this chapter we summarise the key activities and approaches and examine the insights that these emerge.

Best Practice: Establish an Open Data Ecosystem (2016)³² is a late addition, suggesting that this aspect, while acknowledged as key, is perhaps still under-addressed. “More needs to be done” according to the authors, towards, “the establishment of an active Open Data network, an ecosystem to facilitate the uptake of data and information for reuse.” Throughout the articles that follow, the necessity of networks for success, particularly where business, technical, domain and Open Data skills must be combined, is reiterated.

‘New Businesses Around Open Data, Smart Cities and Fiware’ (2015) is an EUPSI Topic Report. Smart Cities are seen as a key enabler of new businesses using data because of the sheer volume of the data they generate, the contemporary challenges presented by urban living and the number of possible applications that can be replicated from one city to another. A key message of this paper speaks to this last point, and the role of common standards and frameworks (here, provided by Fiware) in enabling this. This paper also notes the complexity of value chains for Open Data entrepreneurs in smart city arenas, involving public administrations, the ICT sector and both profit and non-profit users. The combination of Open Data with other data sources also creates the potential for complexity or friction. Another interesting point in this paper is that until now, solutions have been focused on

³² European Data Portal (2016). Best Practice: [Establish an Open Data Ecosystem](#).

societal benefits, but the focus must also expand to engage profit, which they note is required for sustainability of released data.

New Businesses Around Open Data, Smart Cities and Fiware (2015) also focuses on the role of accelerators (in the case of this report: ODINE and the FIWARE suite of accelerators including FINODEX) not only in funding new companies, but in creating the vital ecosystem of mentors, customers, providers that are required to deliver successful Open Data businesses. “Setting Up Open Data Businesses” (2014) is a report on the launch of FINODEX. It is critical of the levels of success of entrepreneurial Open Data businesses to date, stating that the reality is limited apps with no business models, but hailing European-wide Open Data accelerators as the method for correcting that.

The 2016 ‘Best Practice: Support Open Data Start-ups by the Share PSI 2.0 Thematic Network’³³ recommends a very specific kind of accelerator: the academic business accelerator. The authors argue that as universities are experts at knowledge transfer, mentoring and coaching, as well as having the public and private partnerships for accessing funding, they make the ideal home for ‘academic Open Data incubators’. Such incubators have been created in Greece, Spain and the Czech Republic. Notably, they are also a source of potential future entrepreneurs, who can be encouraged and engaged early on and ‘play’ with Open Data in a low risk situation. The 2012 ‘Reuse Innovation Contests’ papers also suggested that a key benefit of such contests was the opportunity to apply their skills in a practical, real world setting.

Another benefit of universities is that they are well-placed to support intercultural communication education. As multiple authors have recognised, scalability across many cities is vital for commercial success, and universities can support this ability to work in various cultures. The authors also mention the ability of universities to engage with the media; like the Reuse Contests paper, they understand the value to entrepreneurs of promotion and visibility.

The EUPSI Topic Report paper “Innovation Contests for Open Data Reuse” (2012) surveyed both app contest participants and organisers. They specify six key benefits of such activities, among them the potential for creation of ‘new businesses or other forms of economic activity’. This paper is an early empirical study of the phenomenon in which shorter contests are less likely to produce long term, sustainable applications, but have a specific role in formulating a community of ‘civic-minded activists’.

The “New Businesses” (2015) paper reflects this when they identify the requirement to engage more with the profit motive for sustainability; this suggests that app contests are perhaps in themselves not a solution, however, combined with the opportunity to engage with ecosystems via accelerators, they can be more successful. This paper also states that while it does not specifically recommend themed contests over open ones, it is useful to compare solutions along specific themes, thus ‘making it easier to select the really good applications’. The authors suggest another successful factor of innovation contests is “side events”. These opportunities to network, attend workshops, hear from experts or participate in ‘Dragons’ Dens’ are considered to assist the participants in staying engaged and producing better results. Although the survey respondents and authors do not use the term, such activities are essentially those that assist in developing an Open Data ecosystem. As in the “New

³³ European Data Portal (2016). Best Practice: [Support Open Data Start Ups](#).

Businesses” (2015) report, scalability is seen as a key criterion for the success of applications and is commonly used as such in judging. The authors of this paper note that as far as most organisers are concerned, reuse innovation contests meet the outcomes they were intended to have, so are considered successes by their own criteria. However, they comment that whereas civically motivated solutions are easy to identify, far fewer commercially successful applications were deemed important by contest organisers. They find this surprising, ‘because of the commercial rhetoric that surrounds the Open Data theme’. In the recommendations, the authors state that to stimulate innovation, it is necessary to facilitate incubation and acceleration. They suggest that the (local) governments running such contests should indeed be the first client or customer of winning solutions to support the entrepreneur by providing the all-important proof of concept. A final insight from the authors concerns prize money. This is often understood to be desirable, but in fact other prizes, such as introductions or promotion of the winning app, are often equally if not more desirable as they may be harder for the individual entrepreneur or start up to facilitate.

ODINE (2015)³⁴, a report on the start-ups accepted to the first two cohorts of the Open Data Incubator for Europe, provides more insight into this. It specifies that in addition to grant funding, the businesses will have access during their six-month incubation to a suite of services including peer-networking, technology, datasets, mentorship, investors and media services. This demonstrates the broad range of support that Open Data start-ups may need to be able to call upon. Three further and later pieces highlight some key aspects of the paper. Lisbon Opens Up³⁵ is a community report on the initiative “Smart Open Lisboa,” a collaboration on mobile apps between the municipality and commercial, third sector and entrepreneurial organisations which reflects some of the insights in the Reuse Contests paper. Nine start-ups were ‘pre-approved’ for support, which included funding but was largely about creating the ecosystem; access to Web Summit was included in the support.

“Apps for Europe: Hackers Oust Mayor of Amsterdam”³⁶ (2013) is a community report on the Apps for Europe hackathon that specifically focuses on attempting to provide the support needed to enable a prototype app to transition to full-fledged entrepreneurial business. In this instance, alongside the hack a ‘Business Lounge’ was provided, which engaged participants with experts, investors and incubators. “Open Energy Days” (2016) is a community report for a three-day energy hack event, aimed at students, entrepreneurs and businesses to develop innovative concepts. In this case the prize for the best solutions is a ‘start-up package’ that includes mentorship in business development and marketing as well as more technical aspects

The continued popularity of such events even four years after the Reuse Contests paper was critical of their commercial impact suggests that to date, no commercially powerful alternative has been identified that can also deliver the existing benefits. However, it is clear that organisers are attempting to bridge this gap through initiatives such as the Business Lounge and supportive prizes.

Telefonica Open Futures is a Use Case, describing how the Spanish telecoms company has created an environment deliberately aimed at stimulating start-ups, in an open innovation approach, which not only supports the entrepreneurs by making available to them otherwise closed aggregated mobile

³⁴ <https://www.europeandataportal.eu/en/highlights/odine-entrepreneurs-awarded-total-€17m>

³⁵ <https://www.europeandataportal.eu/en/news/lisbon-opens>

³⁶ Apps for Europe: [Hackers Oust Mayor of Amsterdam](#) (2013)

data, but also rewards Telefonica when new insights and products are created by start-ups mixing open and other data sources with the Telefonica data. Thus, the incentives work both ways, leading potentially to greater sustainability. In the Smart Open Lisboa initiative data provided to the start-ups was also not just open but was shared from the other partners.

The remaining articles take some more individual routes to supporting Open Data start-ups. “Best practice: Open Data business models and value discipline”³⁷ (2016) is another recommendation from the Share PSI 2.0 Thematic Network. This paper emphasizes the necessity of developing a coherent business model in order to fully exploit the value of Open Data and develop innovative products and services. This essentially summarises Ahmadi’s 2014 work on Open Data Value Disciplines, which categorises five business models in terms of their value discipline or competitive strategy.³⁸ The analysis suggests that the appropriate value disciplines for Open Data businesses are Usefulness, Process Improvement, Performance and Customer Loyalty. However, it does not go further into these aspects, presenting an opportunity for further investigation.

“Commercialising Academic Research through Licensing and Spinouts” (2013)³⁹ is a community report on supporting a different approach to innovation – knowledge transfer from universities to commercial entities. This is not solely focused on Open Data start-ups and this approach is only referenced in this report. It becomes clear that this is not considered to be a prominent role for universities regarding Open Data.

8. Recommendations

Five key recommendations to policy makers and activists emerged from our analysis of these papers.

1. Build the ecosystem

This is crucial in order to support and develop Open Data entrepreneurs. The ecosystem is both the one to be built tightly around the start-up, including domain experts, business experts and technical expertise, but also the wider ecosystem, including access to citizens, journalists and data owners and providers.

- Create a diagram of your open data ecosystem. Where is it strong? Where are the gaps? FINODEX alumnus, Belgian company DataScouts developed such an ecosystem monitoring platform based on Open Data.
- Utilise the ecosystem to engage emerging businesses with novel funding mechanisms and revenue streams
- Develop links between groups working at different stages of start-up development, such as those running hackathons and those running accelerators
- Engage with ecosystem groups to ensure they know how and when they can get involved, especially more difficult to reach groups such as citizens.

³⁷ European Data Portal Best Practice: [Open Data Business Models and Value Disciplines](#)

³⁸Ahmadi, F (2014) [Open Data Business Model Patterns and Value Disciplines](#)

³⁹ EU (2013) [Commercialising Academic Research](#).

2. Engage students and young people from early stages

Educating young people about Open Data value and skills in a low-risk environment may make them more willing to develop the necessary skills for working with (Open) Data and, further down the road, to incentivise them to opt for a job in an (Open) Data driven enterprise and (help) develop Open Data driven solutions.

- Develop ways to introduce young people to the concept and landscape of civic technology
- Engage local schools and universities in open data events
- Promote hack events for school age children
- Make use of free materials such as those available from the European Data Science Academy (edsa-project.eu).

3. Invest in national and regional acceleration and incubation programmes with links to local ecosystems, so that the impact continues after the acceleration period

- Provide virtual accelerators that deliver high impact by providing targeted support;
- Support mentors, trainers and community leaders in the accelerators to replicate successful strategies in other settings
- Assess applicants to incubators on the ability to make good use of networks up and down the value chain as well as on the ability to make good use of funding;
- Don't assume accelerators always mean funding – closer relationships with data owners may be equally valuable.

4. Make the link between business, domain and data skills, as these all need to be present (not necessarily as the 'data unicorn' but in a team of people) for a successful Open Data business

- Organise hackathons and challenges on themed lines to engage with the domain
- Ensure business and 'soft' skills are represented at data-focused events
- Design for sustainability: reward the best idea and the best plan for delivering that idea.

5. Separate entrepreneurial activities from other economic uses of data so that they can be tracked and supported

The Open Data Maturity Model has two dimensions, portal maturity and Open Data readiness.⁴⁰ Open Data readiness is comprised of an assessment of policy, of use and of impact. Under the umbrella of use, governments are asked about their attitudes to supporting re-use. Many countries are aware of data being re-used by their own government, but it would be interesting to also examine their awareness of re-use in the entrepreneurship space. Similarly, the specific contributions of entrepreneurs to both social and economic impact would allow for a more detailed assessment of this dimension. Finally, entrepreneurs can also utilise some of these recommendations to strengthen their own start-ups. Defining and engaging with your open data ecosystem, developing networks up and down the value chain and assessing your team on both data and business skills are key to this.

⁴⁰ Carrara W, Radu, C and Vollers, H (2017) European Data Portal: [Open Data Maturity Landscaping](#).

9. Conclusion

Overall, it becomes clear that there is much to celebrate in terms of Open Data entrepreneurship. Exciting, useful products and services based on Open Data are making themselves part of daily life, and the start-up companies developing them are creating valuable new jobs. There are distinct approaches and programmes focused on supporting this development and entrepreneurs who are engaged with Open Data. It is also clear that Open Data cannot simply be published 'in a void' and that commercial projects are necessary to create Open Data sustainability.

With the exception of some key European accelerators, much of the delivery of programmes is still fairly ad hoc, rather than simply localised. It is vital to reconcile the very local and individual approaches to Open Data ecosystems and start-up support with the aim of creating scalable, multi-city, even multi-country start-ups. As part of this, the initiatives that should be built on are those that recognise that valuable start-ups will only emerge from longer term, multi-faceted programmes. It is crucial also to continue to design accelerators that attract the strongest start-ups, including providing support opportunities beyond funding and allowing for comparability so that the best solutions are identified.

Finally, the Open Data entrepreneurial ecosystem must be treated as something not purely organic, but that must be encouraged and cultivated. These ecosystems need to be understood in the broadest terms, and constantly examined and evaluated for areas of strength and weakness, with resources available to exploit strength and shore up weakness.
