Investing in AI: an overview

Author, Francesco Corea

A Data Science Foundation White Paper

August 2018

----

www.datascience.foundation

Data Science Foundation

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ
Tel: 0161 926 3641 Email: contact@datascience.foundation Web: www.datascience.foundation
Registered in England and Wales 4th June 2015, Registered Number 9624670

Who is best positioned to invest in Artificial Intelligence?

A landscape analysis of AI startups

Itseems to me that the hype about AI makes really difficult for experienced investors to understand where the real value and innovation are. I would like then to humbly try to bring some clarity to what is happening on the investment side of the artificial intelligence industry.

We have seen as in the past the development of AI has been stopped by the absence of funding, and thus studying the current investment market is crucial to identify where AI is going. First of all, it should be clear that investing in AI is extremely cumbersome: the level of technical complexity goes out of the pure commercial scope, and not all the venture capitalists are able to fully comprehend the functional details of machine learning. This is why the figures of the ***"Advisors"*** and ***"Scientist-in-Residence"*** are becoming extremely important nowadays. Those roles would also help in setting the right expectations level, and figuring out what is possible and what is not.

AI investors are also slightly different from other investors: they should have a **deep capital base** (it is still not clear what approach will pay off), and a higher than usual **risk tolerance: investing in AI is a marathon**, and it might take ten years or more to see a real return (if any). The investment so provided should allow companies to survive many potential "AI winters" (business cycles), and pursue a higher degree of R&D even to the detriment of shorter term profits. An additional key element of this equation is the **regulatory environment,** which is still missing and needs to be monitored to act promptly accordingly.

When it comes to **AI hardware or robotic applications** then, few extra points are advised—investors do not have to suffer from the sunk cost fallacy bias, and technical milestones should be clear a priori to track real progress.

All these characteristics are motivated by a series of AI-specific problems: first, as above-mentioned the technical complexity makes often AI startups **black boxes.** Secondly, it is quite hard to show **proof of concepts.** Some narrow AI prototype might be easier to be built, but in general, the difficulty of creating GAI-resembling software and the opaque benefit-costs analysis make hard to attract initial funding—and in my opinion, this is where the governments should intervene in. The concern then about what kind of milestone is deemed investable (revenues, open source communities, etc.) is tangible, and I would suggest considering investable only those companies showing some degree of technical innovation, either actual (MVPs) or potential (academic publications), or with data virtuous cycle (a mixture of unique datasets and users).

On the hardware side instead, other considerations have to be added: they are way more expensive than AI software developments, and victim of **higher obsolescence** and **replacement costs.** Hence, the tradeoff cost/reliability/speed/full control adds a further layer of complexity in the investing game. In particular, it is interesting to notice that if we would be able to work in the robotics space at much lower costs, this would shift completely our risk aversion perception, and it would encourage investors to risk more given the lower cost.

Having identified all these characteristics, we can try to draw a rough profile of companies that might represent (ex-ante) good investment opportunities: an early sign of good potential investment is definitely the **technical expertise of the founders/CEOs.** You should prove to have the right mix of technical understanding, technology exposure, access to a wider network, and vision leadership in order to convince brilliant researcher to work for your AI company. The second point of interest is the **perse and multidisciplinary team:** it does not sound impressing having all the co-founders or research team to come from the same school or previous research lab, but rather quite the opposite. Finally, startups that are **people-centric** are ex-ante more likely to succeed. The ability to create and supporting a developer community, as well as making products that are designed to be easily understandable have more probability to be adopted without frictions.

It is not a coincidence indeed that all the features so far highlighted were observable in early-stage success such as DeepMind. However, as we already emphasized earlier when discussing new business models, DeepMind has not only innovated from a strategic point of view, but it also stressed out the major points of interest for any AI startup. First, **always aim to a general-purpose intelligence:** the value DeepMind is proving to own is the ability to apply their general research in the same way to medical problems or energetic issue. Second, **do not be afraid of public exposure to failure:** challenging Lee Sedol on a live worldwide recording was risky, but the brand reward and resonance obtained from winning vastly overcame the effects from a (potential) public failure.

In order to study more deeply what the AI environment looks like, it has been created a unique customized dataset listing 13,833 startups, tracking down in financial news and SEC filings companies operating in artificial intelligence; machine learning; big data; analytics; robotics; and drones. Data about the company so selected have been filled using mainly Crunchbase[1] dataset, and major incumbents (Table 1) have been excluded.

Company Name	AI Applications
Amazon	Alexa, DSSTNE
Apple	Siri
Asus	Zenbo
Baidu	Duer, PaddlePaddle
Bloomberg	
EMC	
Facebook	Torch, M, FBLearner Flow
Google	TensorFlow, Google Now, SyntaxNet
IBM	SystemML
Infosys	Mana
Microsoft	Azure, Cortana, Xiaoice
Mu Sigma	
Nvidia	
Oracle	
Palantir	
Qualcomm	
Rocket Fuel	
Software AG	
Teradata	
Tesla	
Toyota	

Table 1. Major AI incumbents and most known AI applications developed by each of them.

Ideally, we should spot some common features belonging to all the successful startups operating in the AI space, because the sector of activity largely influences the companies' structure. However, the artificial intelligence landscape is not mature yet, and it might be hard to reach strong conclusions. The novelty of the space can be noticed immediately looking at the companies' age distribution. In Fig. 1, it is clear how the majority of startups were born over the past 5–6 years (the peak has been reached in 2014), certainly because of the reasons we specified earlier and the recent AI wave.

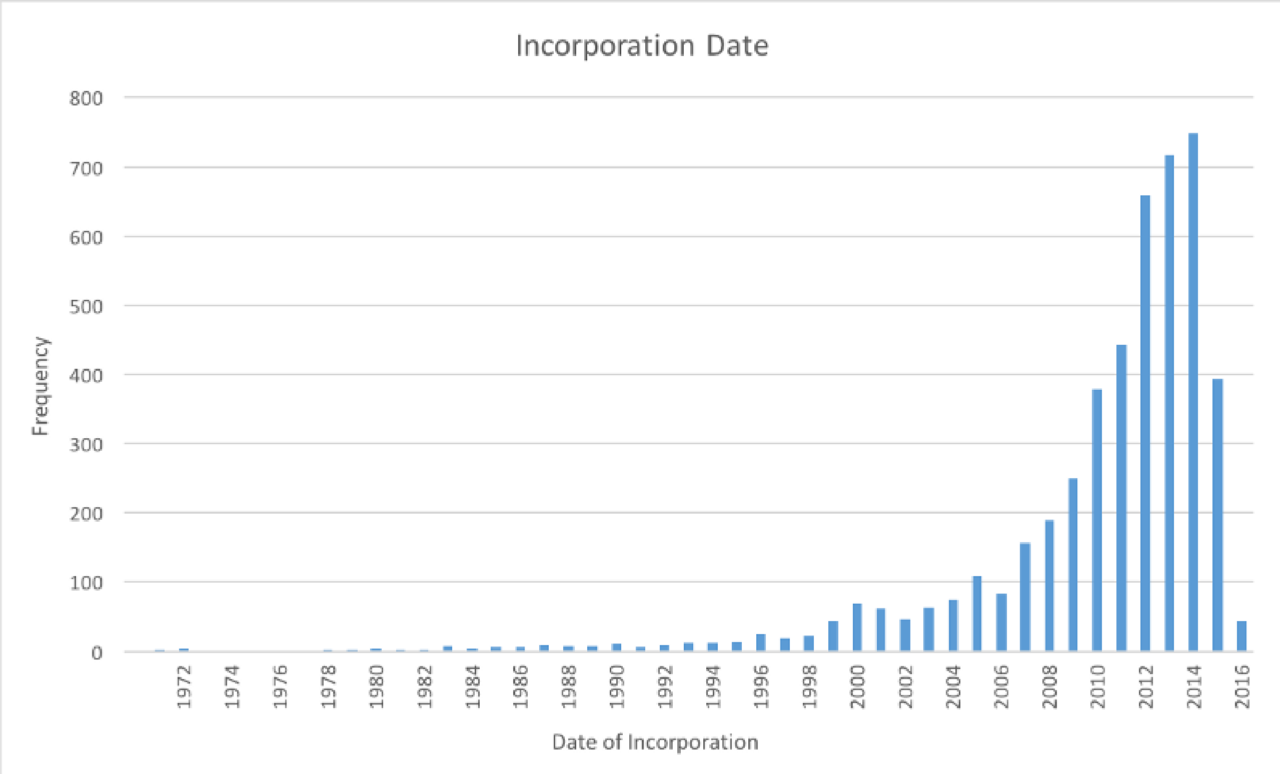

Figure 1. Incorporation date distribution

The geographic concentration of AI companies gives us another insight. Fig. 2 shows that North America plays as expected the most important role in this sector, followed by Europe that accounts for less than a half of the American amount. The Asian ecosystem comes after, and most of the companies operating in Asia are rather in the hardware and robotics businesses. If we look instead at the countries' breakdown, it is not surprising that the USA represents more than 57% of the worldwide AI community. It is relevant though that the English and Indian landscapes are the other important AI clusters (Fig. 3).

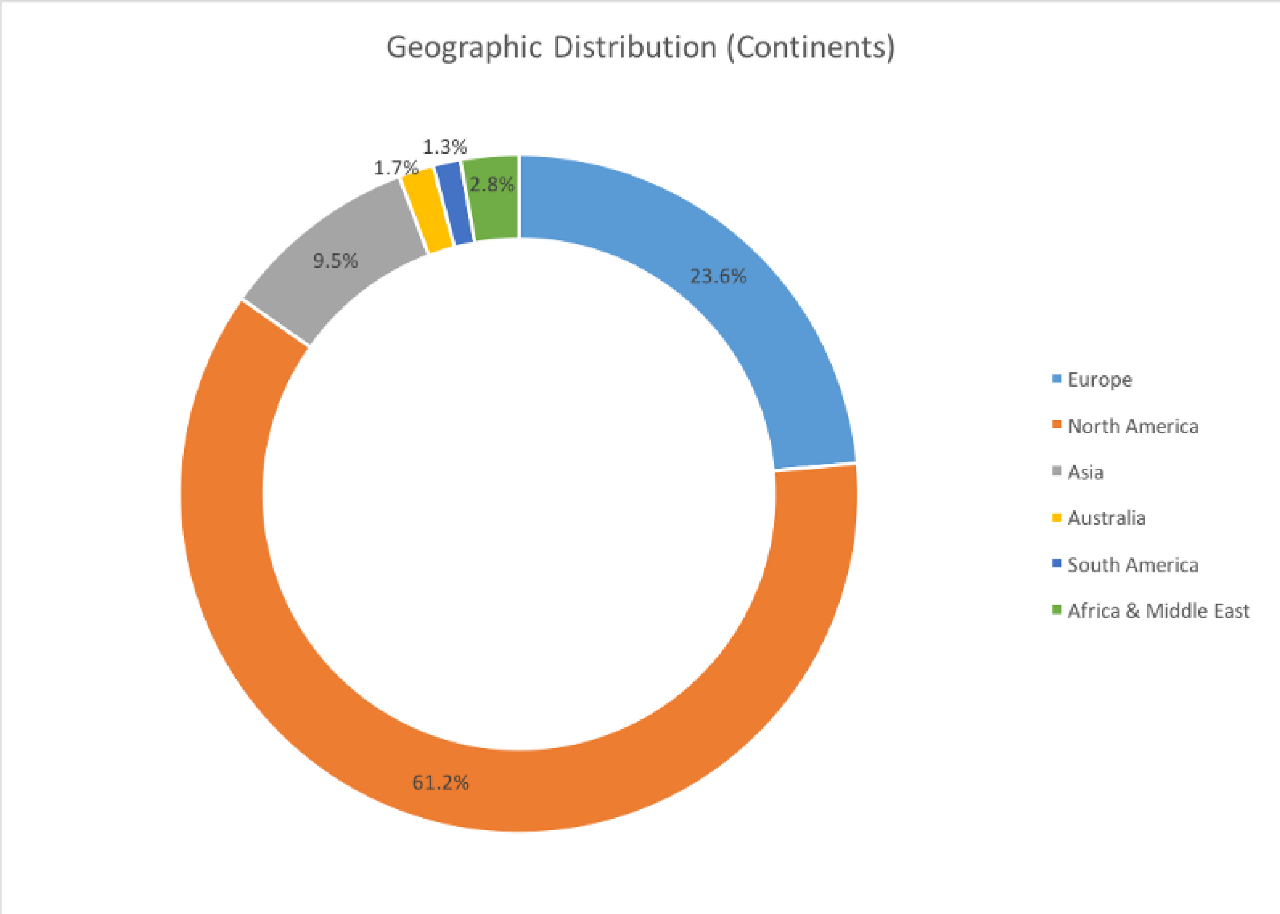

Figure 2. Geographic distribution of AI startups by continent

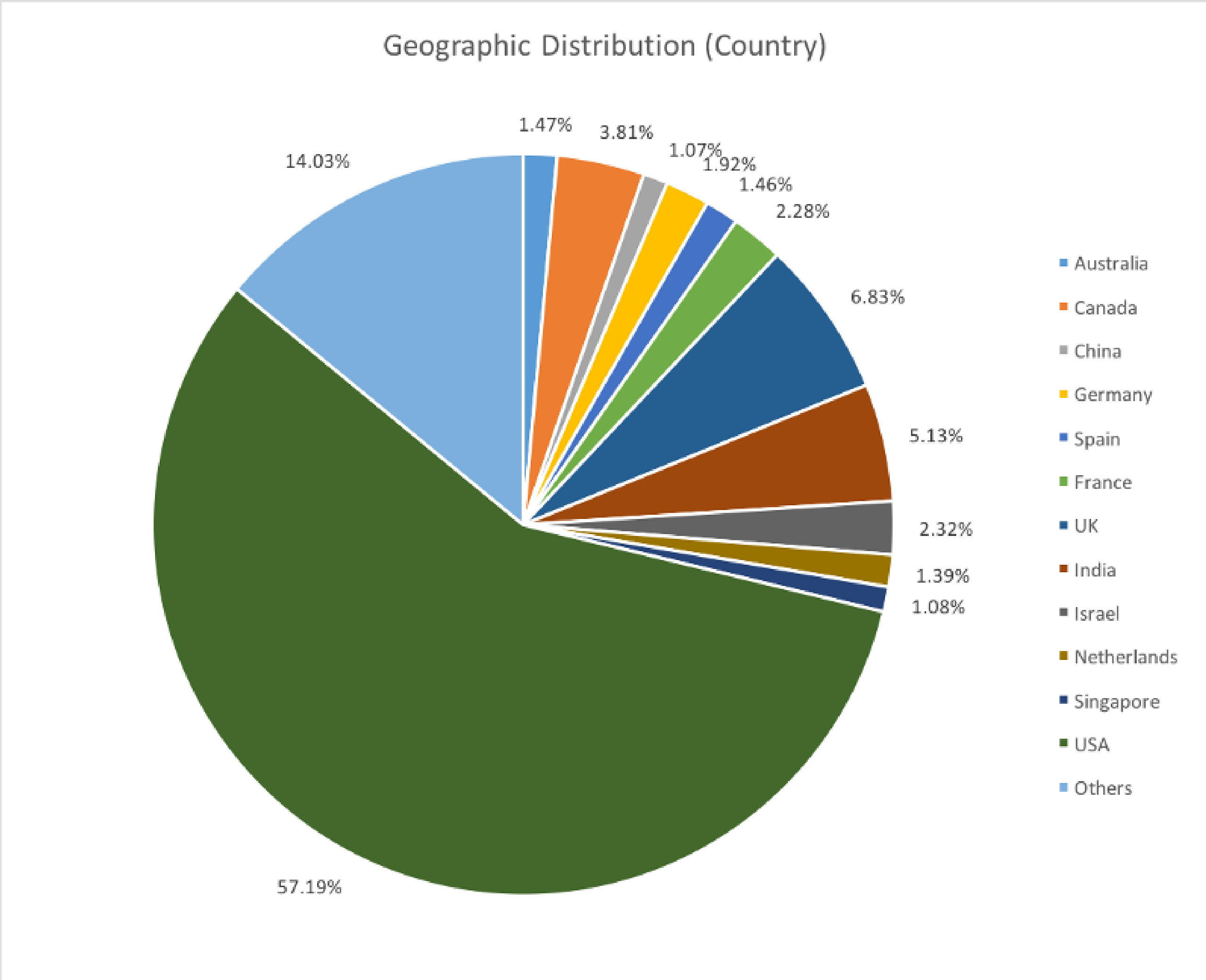

Figure 3. Countries breakdown of AI companies

Digging one further layer (Fig. 4), we notice how San Francisco represents almost one-sixth of the entire market, but other cities such as London or Bangalore are important pieces of the puzzle as well. In fact, if we do not take into account San Francisco and New York that are clearly the two major worldwide startups centers, London and Boston occupy the third and fourth positions. This is not a coincidence, and I believe that these two cities have many similarities. First of all, they are in the middle of strong scientific academic triangles (Harvard, MIT, and Boston University from one hand, and Oxford, Cambridge, Imperial College from the other hand). This fosters the commercialization of academic spin-off and encourages the entrepreneurial culture between students and professors as well. This entrepreneurial wave turned over the past few years into the birth of manifold accelerators and incubators, which are essential for both the ideas generating process and the early development. Finally, the amount of venture funding is critical for the success of the ecosystem. Hence, talent, money and infrastructure are the main reasons of success for London and Boston ecosystem.

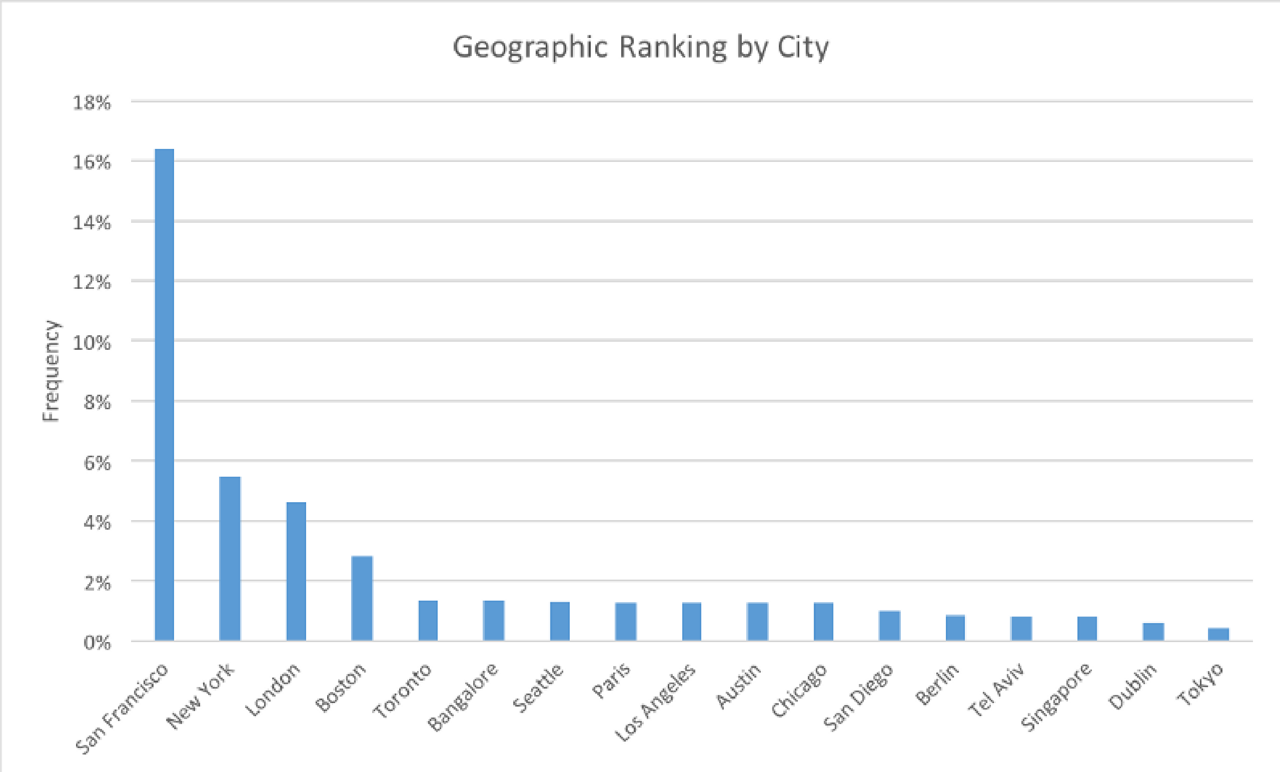

Figure 4. Top 15 cities in the world with AI startups concentrations

There are other two aspects that should be considered in analyzing the environment, namely the financing and the operational sides. From a financial point of view, we can firstly study the time series for the rounds of financing. Looking at Fig.5, which shows the breakdown of the round of financings by year over the past 16 years, we can notice how earlier stages of financing are decreasing on percentage in the last five years and the funding is redistributed to later stage. Despite that, Fig. 6 illustrates that the total amount of funding has drastically increased in the last 2–3 years. Those two insights suggest that usually, AI startups ask for lower rounds of financings (Fig. 7), and often they even barely reach rounds C or higher, either because they are not able to deliver what has been promised or because they are acquired by big players.

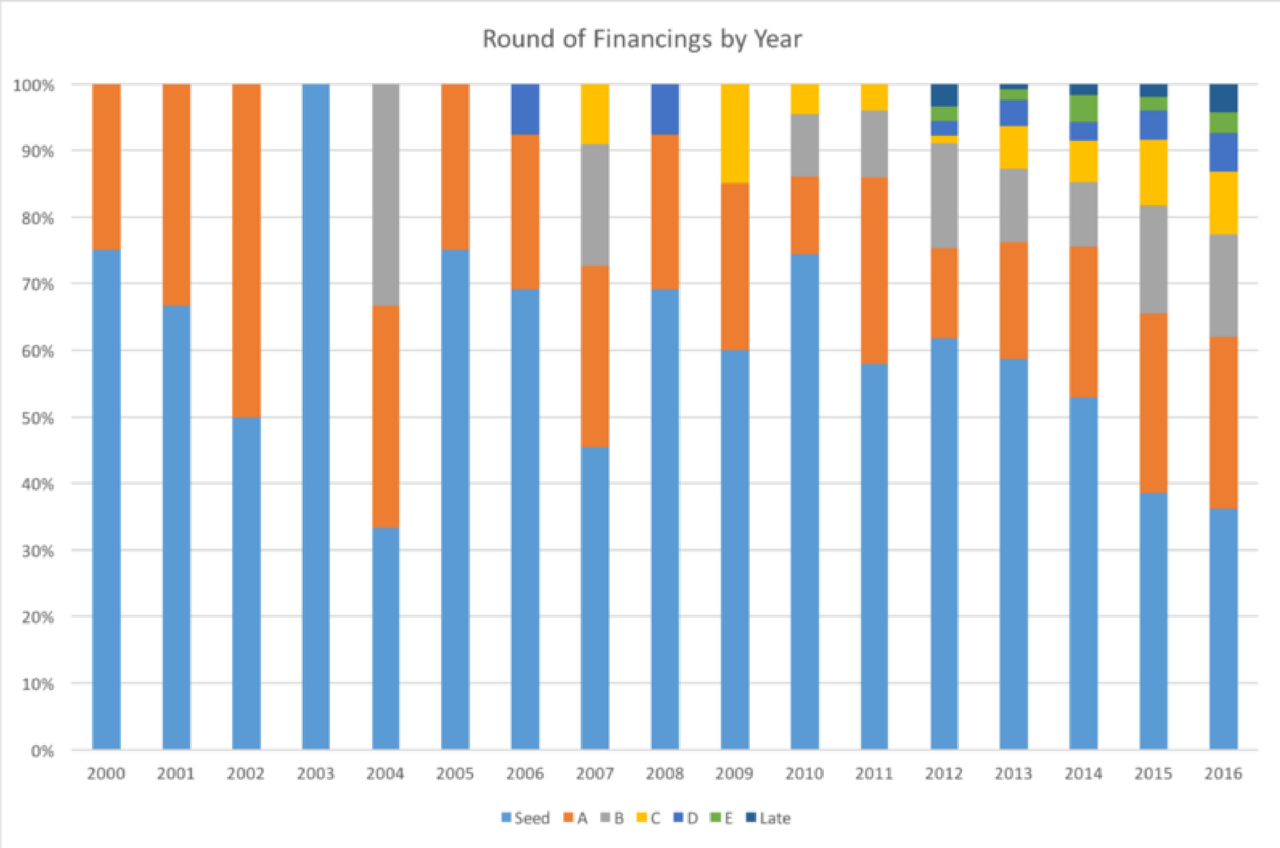

Figure 5. Round of financing breakdown by year for the period 2000–2016

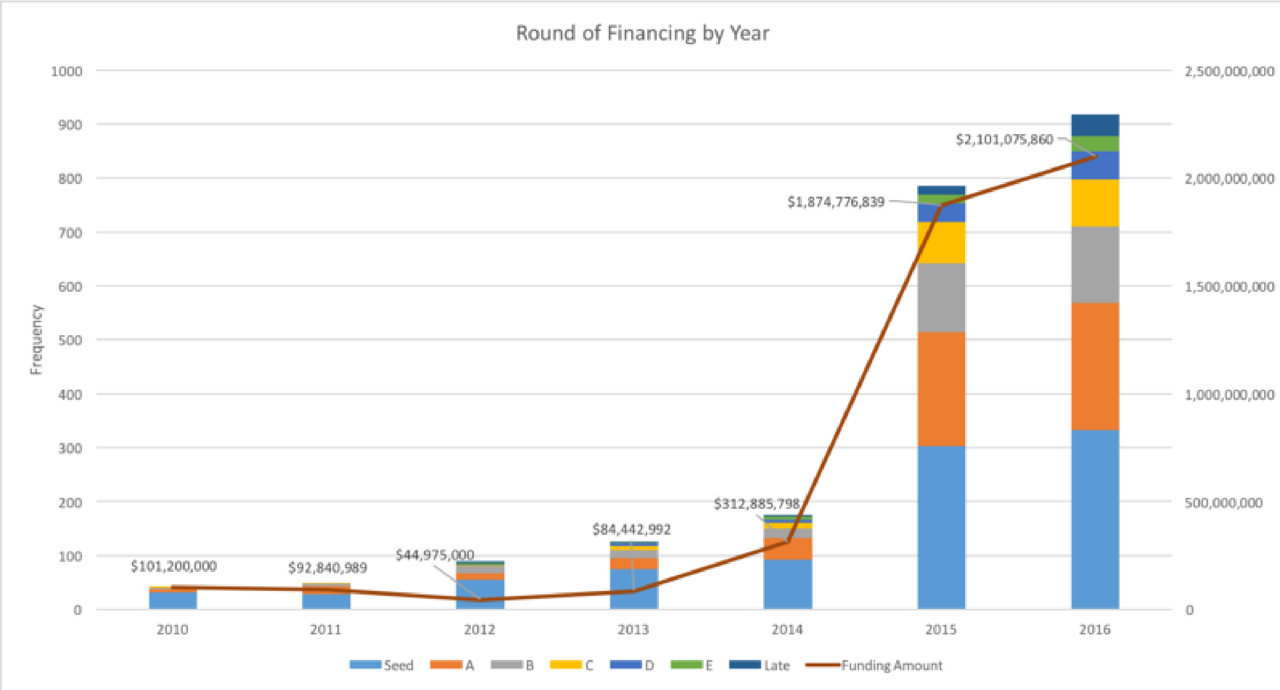

Figure 6. Round of financing breakdown by year for the period 2010–2016 (main axes); total amount of funding by year (secondary axis)

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ
Tel: 0161 926 3641 Email: contact@datascience.foundation Web: www.datascience.foundation

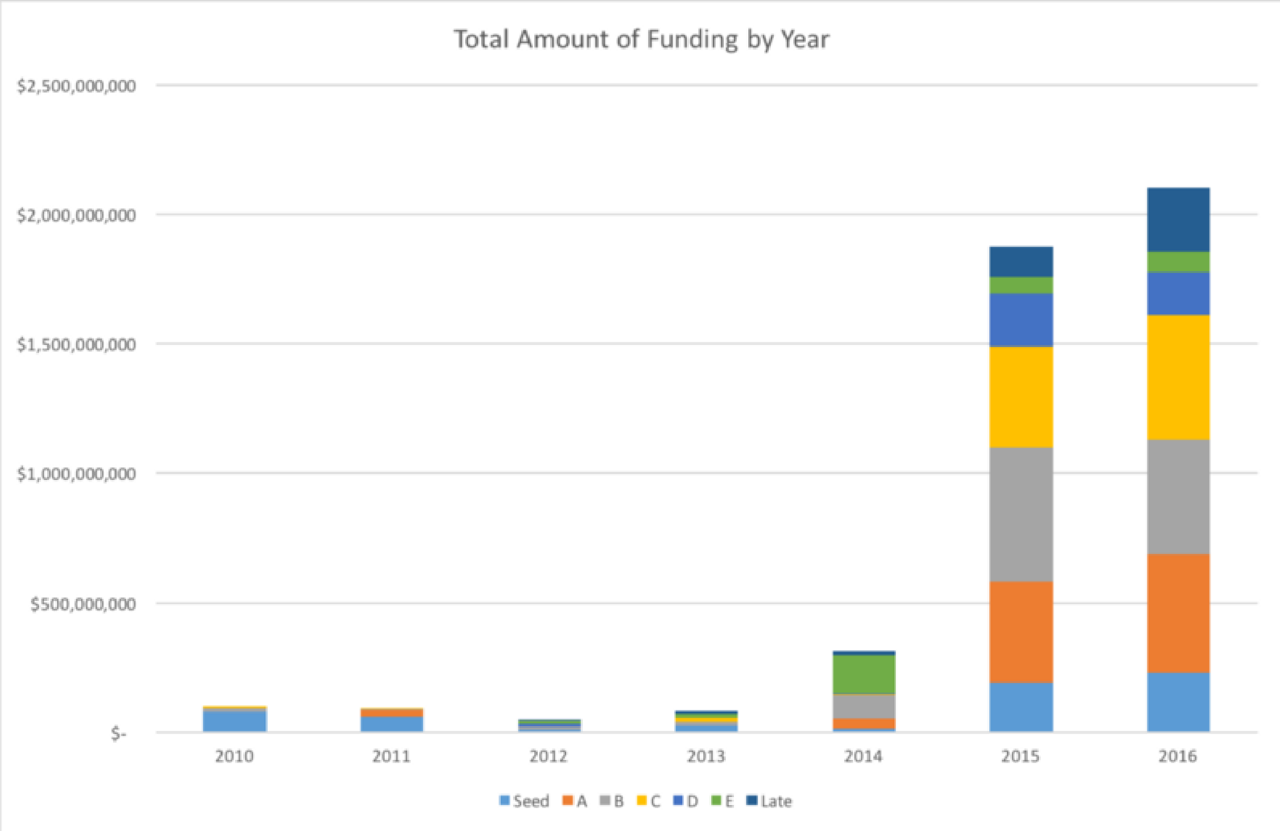

Figure 7. Total amount of funding by year for the period 2010–2016, analyzed by stage of financing

We can confirm this intuition looking at the number and types of exits for AI companies. Fig. 8 shows that a quite high number of startups exits being acquired, while a lower number raises funds in the public market.

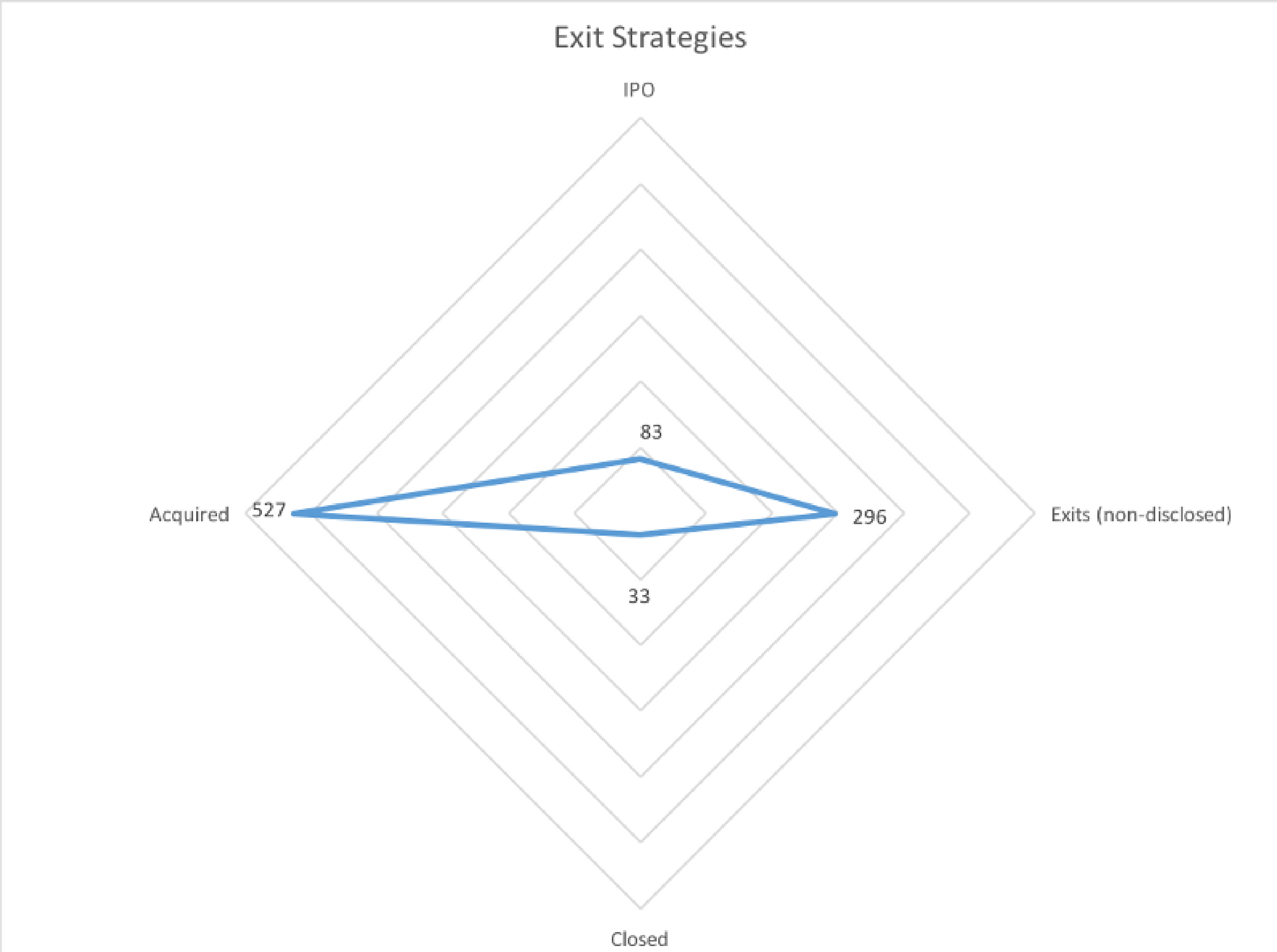

Figure 8. Exits strategies distribution

The operational point of view can be instead faced through studying concrete variables of growth. First of all, we can notice in Fig. 9that the majority of AI startups are composed by up to ten people, and in some cases, they reach forty or even fifty elements. Even if the employees average is around 69 people, the median is rather around 8 persons per company—that on the occasion of exits often means a range valuation per employee of $2.5—$10 million. Furthermore, usually the first ten hired are mostly engineers or technical people, and just after the first round of financing further horizontal layers are added.

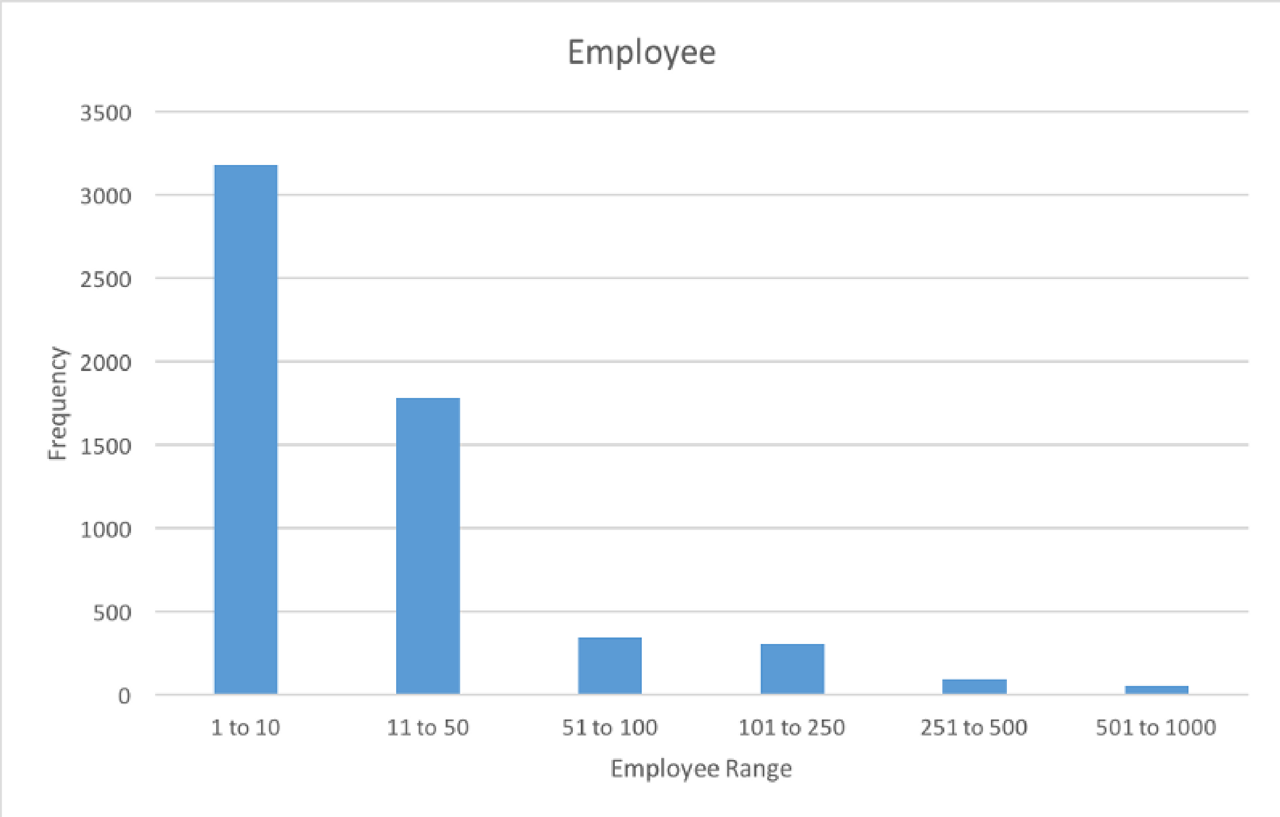

Figure 9. Frequency of employee numbers per company

As growth measure, we looked instead at the employee growth on a monthly and semester basis (Fig. 10). Even if on a monthly basis it is quite normal to oscillate the total number of employees between -10% and +20%, on a longer time period many startups reach exponential growth rates close to 40%-50%.

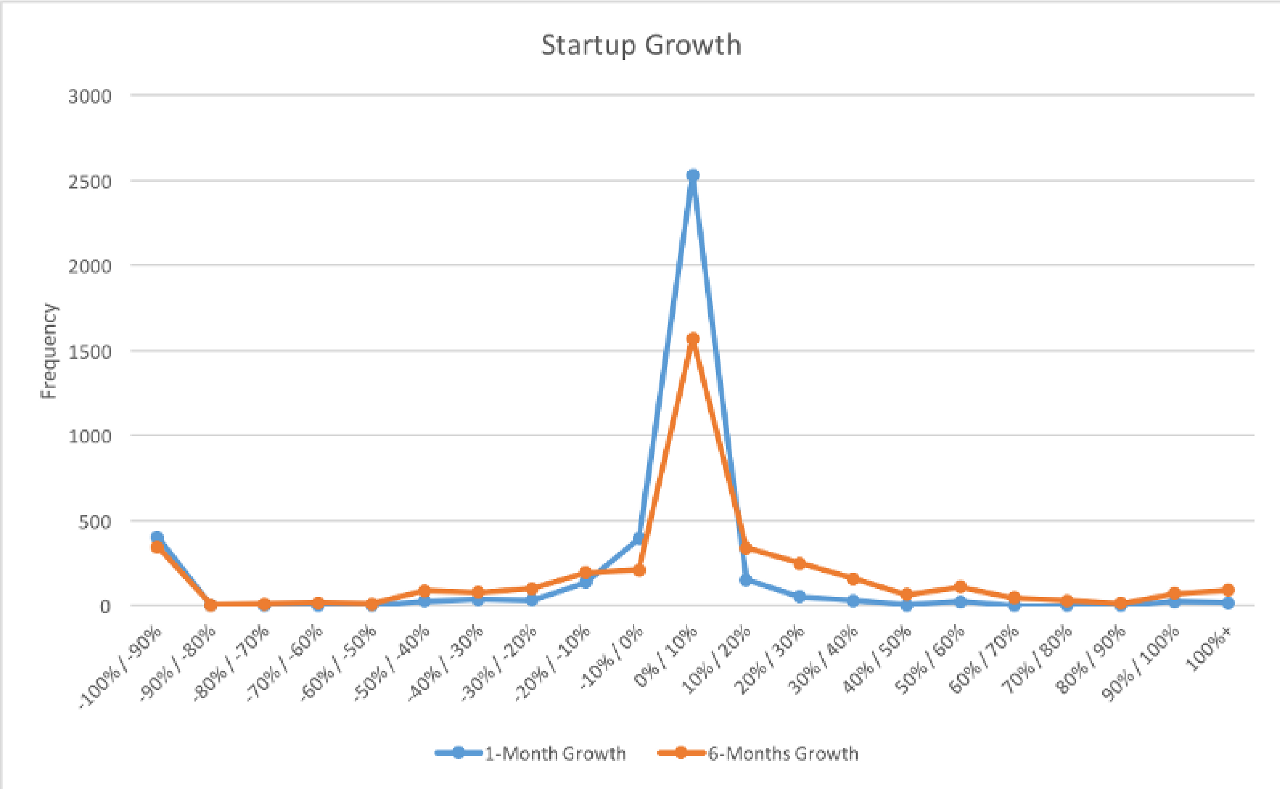

Figure 10. Startup growth rate measure by 1-month and 6-months employees increase

The second proxy of growth is the impact the startups have through social media channels such as Twitter, Facebook, and Linkedin (Fig. 11). Month-by-month social media exposure grows between -10% and +20%, and this validates the idea that AI is gradually being socially acknowledged and used.

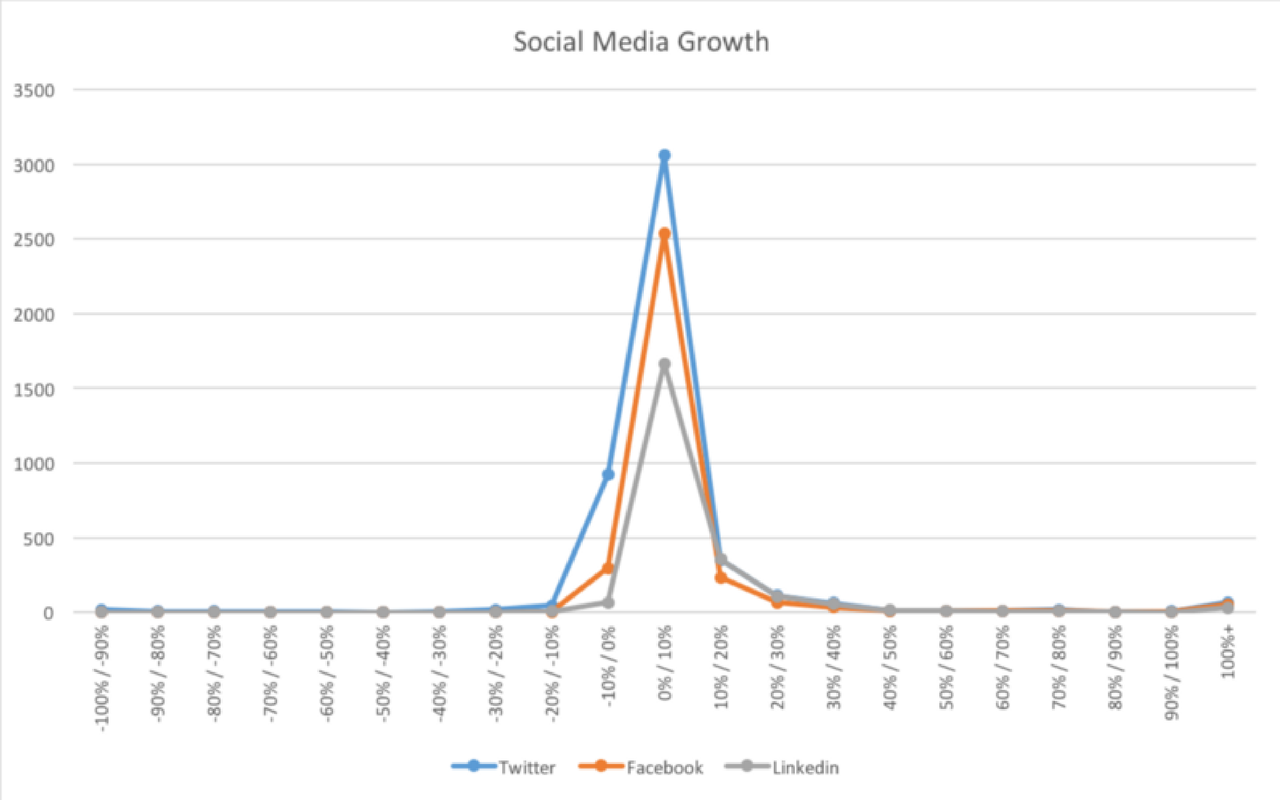

Figure 11. Social media growth on monthly basis

These are only a few common features we can find in AI startups, but there are for sure characteristics we are missing. A point of concern, for instance, is the **average equity stake** investors ask for funding AI startups. My hypothesis is that the average equity required is lower than what is asked in other sectors and that control is less relevant for those investments. The insight behind it is that the technical difficulty in understanding the product makes the venture capitalists contribution less effective from a product standpoint, and it is just limited to UX and market strategies.

The second concern is the **funding resiliency to the business cycle** and to the level of optimism of the mass about this technology. A negative phase might, in fact, pull back all the investments made, because as we explained before it is often hard to see profits in the short term. This might negatively impact the sector overall since the whole AI environment has been pushed for venture funding and corporate acquisition (see table below)—interestingly enough, all big technology players performed poorly in 2015, and this has been already noticed in their tighter acquisition strategies.

The final point to highlight is the way in which **the AI ecosystem** is developing. We are observing a predetermined pollination process: the startup creates an MVP, and maybe launch a first version on the market. It needs for liquidity, and usually get one or two rounds of funding—and thus grows, and hires as a consequence. As soon as the startup is getting some real traction, employees start leaving and they create their own things, encouraged by the idea of being operationally backed and financially supported by VCs. We are not able to conclude with a positive or negative feedback on this mechanism, but it looks at a first glance a bit unstable to lay the foundations for the research of a general artificial intelligence.

Notes

[1] http://www.crunchbase.com. I have obtained a Crunchbase License that allowed me to complete the dataset with relevant missing information of several companies.

This article is an excerpt from my book “Artificial Intelligence and Exponential Technologies” (2016).

About the Data Science Foundation

The Data Science Foundation is a professional body representing the interests of the Data Science Industry. Its membership consists of suppliers who offer a range of big data analytical and technical services and companies and individuals with an interest in the commercial advantages that can be gained from big data. The organisation aims to raise the profile of this developing industry, to educate people about the benefits of knowledge based decision making and to encourage firms to start using big data techniques.

Contact Data Science Foundation

Email:contact@datascience.foundation
Telephone: 0161 926 3641
Atlantic Business Centre
Atlantic Street
Altrincham
WA14 5NQ
web: www.datascience.foundation