

Big Data, Small Data, and Corporate Innovation

Recognise, implement & maximise

The Colinked Corporate Innovation Club is a community of over 120 international and global corporate brands as well as some government departments and NGO's. All the members of the community share a common responsibility and ownership for their organisation's engagement with open and collaborative innovation or digital transformation strategies. The community is by invitation only and has grown now to be the largest community of Corporate innovators in the world. The community is completely free to join and meetup six times a year to discuss a topic or theme which has relevance to their innovation vision, strategies and execution. We also partner with other community groups on trade missions and executive roundtables for our members to attend. The meetings are always hosted by a Corporate member of the Corporate Innovation Club.

Since January 2019 we have also created an online collaborative innovation platform for Corporates, startups and Investors from around the world to connect with each other and the innovation programmes the Corporates run. This online community is also completely free to join and takes minutes to register as a member. Find more details on the platform and register to participate at www.colinked.io.

Meeting Theme

March 2019's Corporate Innovation Club meeting looked at the role data can play in Corporate innovation; particularly in terms of interpreted datasets.

Data interpretation of the kind discussed at the meeting effectively emerged in some form in the 1980s as a practical tool with the arrival of widespread digital media and storage. By the 1990s it had become a valuable part of guiding various disciplines within forward-thinking Corporations, and through the opening decade of the new millennium 'big data' specifically become something of an adored buzzword. As with virtual reality, artificial intelligence and blockchain, hype and overenthusiasm from press, marketers and those seeking investment arguably muddied the waters with regard to understanding of 'big data' and its potential.

More recently data interpretation has settled into something of a more nuanced – if rapidly evolving – field. As such, the Corporate Innovation Club’s members gathered to learn from experts, brainstorm ideas, share insights and collectively gather a clearer understanding of data interpretation as it is today, and the potential therein to bolster the efforts of Corporate innovation.



Key Takeaways

Beyond ‘big data’

Too often the term ‘big data’ is used to frame conversations about using any datasets to guide or inform Corporate innovation. While ‘big data’ emerged as a valid term with a specific meaning, in common usage it has become a general catch-all used when describing datasets. The problem there, however, is that ‘big data’ as a term fails to recognise the diverse spectrum of datasets that can be useful in Corporate innovation and beyond. Simply put, ‘big data’ refers to certain kinds of dataset; not all.

It was put to the Corporate Innovation Club that datasets big, small – and even those that are simultaneously big and small, or neither – can all have a value in an innovation context. When it comes to data, there is not necessarily a correlation between size and value. Those wishing to benefit from data that can inform and support innovation initiatives may have to communicate to other departments and C-suite that meaningful use of data is about more than simply having access to ‘lots of data’. Indeed, small data can often be more relevant and powerful than big data, particularly in identifying causation.

Rather than size of dataset being important, the relevance and application of a given dataset will define its value to a given Corporate innovation initiative.

“Big Data as a term fails to recognise the diverse spectrum of datasets that can be useful in Corporate innovation and beyond”

A matter of scale. Defining the difference between ‘big data’ and ‘small data’

While the full spectrum of types of dataset available is wildly diverse, it can be useful to consider the broad distinction between big and small data.

Broadly speaking, ‘big data’ refers to datasets so enormous interpretation by humans alone is essentially impossible. These datasets are potentially useful, however, because - through the use of technology - correlations can be found that can inform or support a myriad of different endeavours. The likes of automation, machine learning and artificial intelligence can be powerful allies in interpreting and making use of big data.

Small data, meanwhile, refers to smaller data sets, enabling observers to gain greater understanding of causation.

Consider a global logistics business. The movement of every item and vehicle, the inventory of every warehouse, each employee’s interactions with internal systems, the usage of thousands of engines and machines, and the flow of the supply chain may collectively kick out data that makes up a ‘big data’ dataset. The automatically updated schedule of a single warehouse manager, meanwhile, offers a small data dataset that is easily digestible and understood by a single human. Small data can often be part of big data in this way. Indeed, many big data sets can be considered as collections of interlinked or overlapping small data.

“Indeed, many big data sets can be considered as collections of interlinked or overlapping small data.”

While big data in the case of a global logistics operation may help identify trends or problems within that business, the small dataset may more meaningfully explain the cause.

As such, using both big data and small data in tandem can be remarkably powerful. While big data interpretation can identify correlations – perhaps between human behaviours and use of a technology – small data can help identify the cause of such correlations.

Practical examples of applications of data interpretation

Through the ‘digitisation of life’ data is generated in huge volumes every day. The way we use our phones, the ways customers use our products and the way a workforce interacts within a Corporation, for example, are all actions that can spawn data in a digital world. As such, interpreting datasets may be used to:

- Increase efficiency within a Corporation
- Monitor and guide the success and impact of a longform innovation initiative
- Refine the user experience of a product or service; internally or externally
- Predict and pre-empt problems customers or clients may experience
- Develop and refine a product or technology being implemented through partnership with a startup or scaleup



Data interpretation is a highly specialised field

On a number of occasions throughout the gathering of the Corporate Innovation Club's members, it was noted that interpreting data in a way that is meaningful, accurate and implementable requires a great deal of experience and specialisation. That is the reason many larger corporations already using data to guide their business beyond Corporate innovation house internal data labs.

The lesson for those in Corporate innovation? To make the most of interpreting data, be open-minded to partnering with external experts, collaborating with relevant startups and scaleups, and working with your own corporation's data labs.

Quantifying data's impact; standardising its mechanisms

Those in Corporate innovation will be very familiar with the challenge and opportunity of being able to measure and communicate the impact of a given initiative. A comparable challenge exists in the realm of interpreting data. Increasingly, there is a need to create mechanisms to standardise and share meaningful data. Presently datasets large and small take many different forms and structures, and are collated and interpreted in many different ways. To a degree there is a need for that diversity; not all datasets can have the same foundation, or will serve the same purpose. However, there was general agreement that an effort to standardise and unify the datasets used across different Corporations and disciplines would be broadly helpful. That may mean rival and competitor Corporations communicating and collaborating. The more we share lessons learned from the gathering, collating and interpreting data, the more useful data can be across the spectrum of roles it serves.

“The more we share lessons learned from the gathering, collating and interpreting data, the more useful data can be across the spectrum of roles it serves.”

Key considerations when gathering internal data

For a Corporate innovation team to get the most out of working with an internally generated dataset, the experts in attendance offered a fundamental guide to approaching the task.

- There is, arguably, no such thing as discovery without a hypothesis. Start by asking 'why are we collecting this data?'. What does the team hope to achieve? With that established it is time to target specific data generated within a corporation.
- Do not assume all data available – even considering targeted data – is useful or meaningful. Use expertise to highlight what can serve as valuable data.
- Having 'too much data' is a common problem. People within Corporations can feel inundated by receiving too many reports containing overwhelming amounts of data. Early on in a data driven project or initiative, think about how any final results or reports can be delivered in a way that is simple, elegant and readily consumed or understood.



Data interpretation is a highly specialised field

It was suggested that sometimes Corporations can become too focused on internal data, or data that plainly correlates to a business' core proposition or offering. The example was given of a hypothetical agricultural tech company's innovation initiative. While that company may be tempted to focus on data generated by its own agricultural endeavours, pulled directly from its partner farms, data from a satellite company detailing landscapes and environments broadly may be more useful to achieve a given goal.

Computing power and big data

As we rely on computers to read and interpret true big data, we are ultimately beholden to existing computing power when it comes to harnessing the big data opportunity. As such, the rise of quantum computing is often presented as a key opportunity with regard to maximising the potential of data interpretation. Understanding quantum computing is incredibly difficult. The approach uses remarkably complex approaches taken from quantum mechanics within the field of physics.

“As we rely on computers to read and interpret true big data, we are ultimately beholden to existing computing power when it comes to harnessing the big data opportunity.”

What is important for those in Corporate innovation to understand is the impact of quantum computing, rather than the inner workings of these deliberately unstable machines. It was posited at the Club meeting that if we consider current data interpretation computers to be 'walking', realised quantum computing would be equivalent to 'space travel' on the same scale. Quantum computing could be a workable reality within as little as three years. At that point we could take all the data ever encrypted in the past thirty years and decrypt it in mere milliseconds. Clearly, quantum computers will easily and rapidly handle multiple highly complex big data datasets. To harness that opportunity Corporate innovators will need to look at using multiple datasets. As such collaboration, data sharing and standardisation as detailed earlier in this reports will become all the more important.

Quantum computers will also be capable of mining sparse data sets millions of times over, going far deeper and pulling far more meaningful, actionable information than currently possible.

For Further Discussion/Consideration

- As data becomes more valuable to Corporations, how important will data ownership become? Does it belong to the person that created the data, or those that gathered it? If individuals want to own and control the data they generate, could data even serve as a new form of currency?

- How will we respond to any emergence of fraudulent data, and even weaponised data?
- What role must regulation and ethics committees play in guiding the increased use of interpreted datasets?

white • space

Whitespace is focused on helping high growth technology startups build a strong and repeatable business as quickly as possible. We do this by learning from startups and Corporates that we closely partner with and the combined experience of our founders. We identify the common areas where startups struggle or make mistakes and we build Intelligent Cloud technology solutions to break down barriers to entry for the Founders around business planning, investment raises and professional services that are needed to help mature and scale their businesses.



cic@white.space



Whitespace, 3rd Floor, 58 City
Road, London EC1Y 2AL



www.white.space