Adding data to your IP monetisation playbook: how to make sure your company is ready

Alongside other intangible assets, companies are deriving a growing amount of value from the accumulation of large volumes of data. But as with patents, valuing and then monetising those newer assets throws up various challenges

By Efrat Kasznik

The advertising world has come a long way since the 1960s, when the business model could be captured in the words of Don Draper of *Mad Men* : "People tell you who they are, but we ignore it because we want them to be who we want them to be." Fast forward to 2020 and not only are people's personal preferences not being ignored, they are being carefully recorded and collected through search engines and other online applications, analysed by powerful algorithms and then made available to brands that wish to target their advertising and thus increase the odds of a product purchase.

As Google – global leader in advertising revenues – states: "The goal of our advertising products is to deliver relevant ads at just the right time and to give people useful commercial information." The driving force behind the transformation of the advertising business in the past 50 years is the availability of large amounts of data related to consumers' buying decisions, aided by technologies that enable its digitisation, collection, storage and analysis of to greatly improve the return on advertising revenues.

In 2017, *The Economist* declared that "the world's most valuable resource is no longer oil, but data". The 21st century economy relies on data, which is a new type of intangible asset that can be viewed as the digital asset that is fuelling companies in all sectors, from banking to manufacturing to biotech. 'Data' can be interpreted to cover a wide variety of compilations of information, but this article will refer to digital, readable, machine-accessible formats.

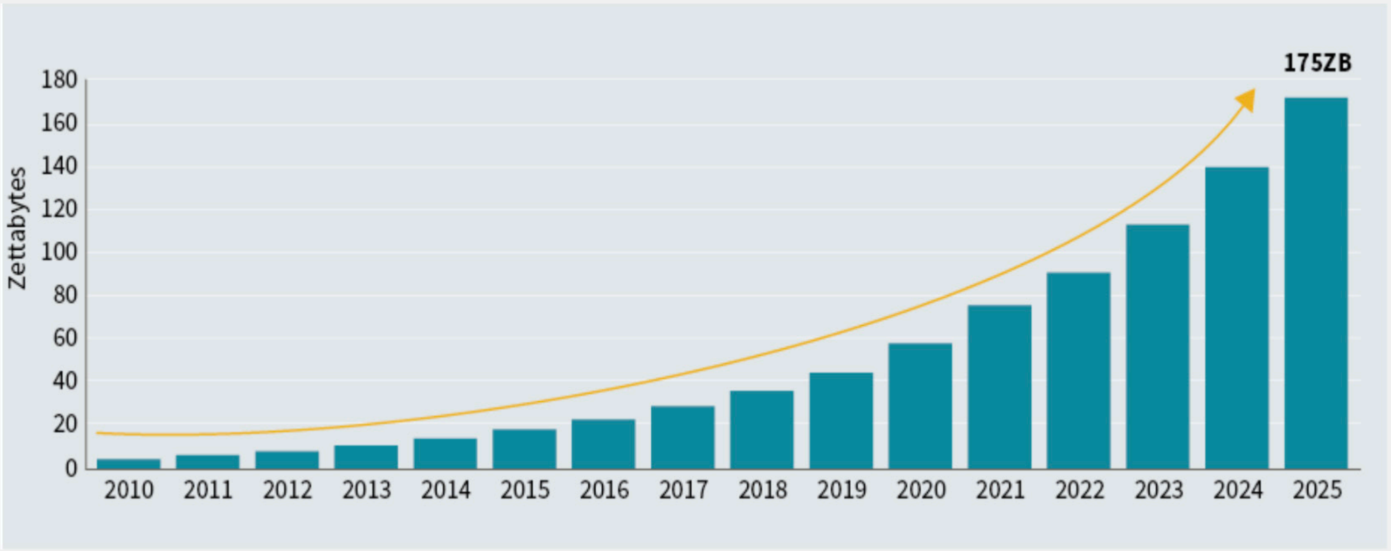
In Data Age 2025 (May 2020), research firm IDC defined three primary locations where data is created and located:

- the core (traditional and cloud data centres);
- the edge (enterprise infrastructure and branch offices); and
- the endpoints (PCs, smart phones and Internet of Things (IoT) devices).

IDC predicts that the global datasphere (defined as all data created, replicated or stored in the above three locations) will grow almost four times, from 45 zettabytes (ZB) in 2019 to 175 ZB by 2025 (1 ZB = 1 trillion gigabytes), as seen in Figure 1. Companies serving the Cloud are seeing skyrocketing growth and valuations; in September, Snowflake – a cloud data warehouse company (an area of services that did not exist a decade ago) – enjoyed the largest initial public offering of a software company ever, raising about $3 billion

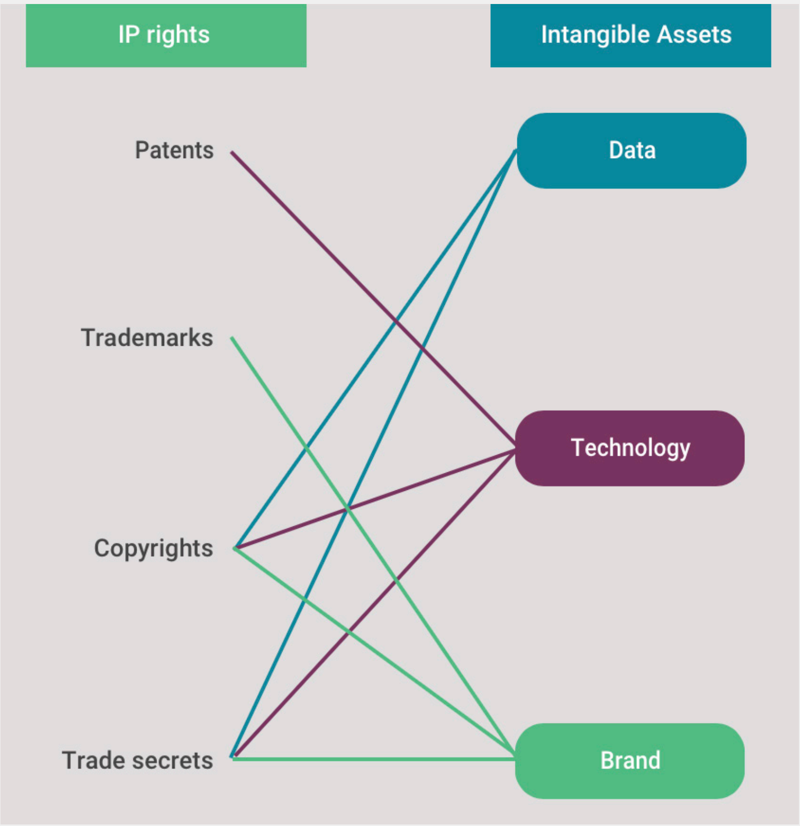
Picture: spainter_vfx/
Shutterstock.com

FIGURE 1. Annual size of the global datasphere



Source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, May 2020

FIGURE 2. Mapping intangible assets to IP rights



Source: Foresight Valuation Group

(based on opening day prices) at a valuation of more than $70 billion.

The digitisation of information facilitates the query and analysis of large quantities of data, which enable new business models for monetisation. Networking and cloud technologies allow the transfer and storage of large amounts of data with easy access for analysis, and AI is changing the way in which data is interpreted for business decisions. That being said, the emergence of the Internet of Things (IoT) and other decentralised ecosystems for data collection through networks of sensors and devices creates challenges around ownership, privacy and protection. A new paradigm viewing the enterprise as the 'steward of data' imposes obligations and regulations related to the prudent way of collecting and leveraging this intangible.

This article explores the monetary potential of data assets and their contribution to overall corporate value. Since data is not covered by patents, its monetisation is not necessarily driven by exclusionary rights, but rather by its ability to drive consumer buying decisions or improve business processes. We will identify the monetisation challenges while exploring the emerging business models for data exploitation through some prominent examples in the market today. Many of these models represent a paradigm shift from how other key intangibles (eg, patents) are being valued and monetised.

Aligning IP protection with data assets

The key to developing a strong IP position starts with investing in the creation of the intangible assets that bring the most value to the corporation and then securing the appropriate domestic and global IP rights to protect these assets. Figure 2 demonstrates the various correlations between intangible assets and the IP protection that they are usually associated with in a typical operating company.

Here, intangible assets have been divided into three categories. It should be noted that the term 'intangible assets' is normally used in the context of financial reporting to denote assets that are neither tangible nor financial and – while the accounting definition includes different allocations for those assets – below is a three-category structure for simplicity:

- Technology – this covers assets encompassing the company's developed technology and/or software algorithms that directly result from the company's

R&D and engineering efforts and underlie the company's products and services.

- Brand – this includes assets related to the company's marketing operations, which help differentiate the company's products and services in the market and enhance its reputation and connection with customers.
- Data – this includes all digital intangibles that relate to the company's operations, production, customers, competitors or any type of information that is digitally collected and maintained by the company and brings it value.

Each of these groups of intangible assets has been linked with one or more commonly associated types of IP protection (the mapping presented here is not intended to be an accurate legal representation of all possible types of protection across all jurisdictions, but rather an operating model that can help to guide a business strategy around IP protection). Most technology companies hold many – if not all – of these three groups of intangible assets, along with their associated types of IP right, which collectively comprise their IP portfolio. There is a difference across industries and products as to the relative weight associated with the value of each type of intangible asset. For life sciences companies, for example, the technology bucket will usually be the most valuable one as a result of significant R&D spending, which leads to heavy reliance on patent protection. For consumer goods companies, where significant resources are allocated to marketing spending to create a strong brand, trademark protection will capture a relatively higher portion of the IP portfolio.

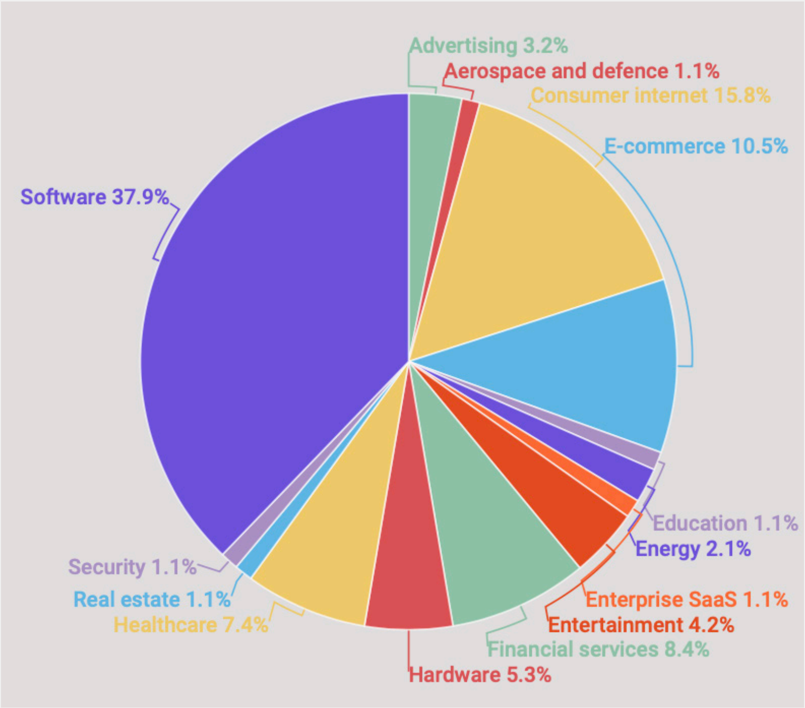
One of the most interesting observations on the map linking intangible assets with IP protection is actually the missing link between data and patent protection. Unlike the technology bucket, data assets are not protected by patents. The main IP protection afforded to data assets is trade secret protection, which is generally implemented through mechanisms such as strict authentication measures around the access to data, cybersecurity defences warding off cyber-attacks and strongly-worded legal contracts governing data access from both inside and outside the organisation. This is an important insight to keep in mind as we continue exploring how data assets contribute to corporate value.

Data assets propel corporate value

Data assets have been growing in significance as one of the key drivers of corporate value, particularly in software companies where digital information is more easily generated though users. Further, their importance has been growing gradually in many other types of company along with the proliferation of IoT ecosystems in many industries. However, the lack of patent protection for data assets may create an issue for some companies where IP protection appears to be misaligned with corporate value. As patents are the most observable form of IP protection for tech companies, the fact that data assets are considered valuable and patents are not available as a mechanism may lead to the misinterpretation of lack of patents as lack of IP value.

This is particularly common in software companies, as seen in valuation studies of unicorn companies (pre-exit start-ups with valuations exceeding $1 billion). A study of the CrunchBase leaderboard of 95 US unicorns (Foresight, 2015) revealed a sample consisting of the vast majority (more than 65%) of software companies, across the advertising, software, consumer internet, e-commerce, enterprise Software-as-a-Service (SaaS) and security categories (see Figure 3). Overall, the study concluded that 62% of US unicorns had 10 or fewer (issued and pending) US patent assets to their name (see Figure 4); these companies accounted for more than $157 billion in collective valuation and $25 billion in combined funding. Even though several of these unicorns (eg, Uber) set out to build strong patent portfolios after achieving unicorn status (through patent acquisitions and organic filing), having patents was certainly not a prerequisite to achieving unicorn status. Understanding the key role that data assets play in the valuations of these unicorns, as well as the relationship between data and patent protection, can explain some of this perceived mismatch between unicorns' corporate value and their patent positions.

FIGURE 3. US unicorn distribution by industry



Source: Foresight Valuation Group

Software unicorn valuations are an example of the data-centric valuations that started showing up in the late 1990s in the early days of the Internet. Large funding rounds and acquisitions of pre-revenue companies have gradually become more common, particularly when it comes to software companies in business-to-consumer (B2C) verticals (eg, social media).

With the advent of smartphones in the mid-to-late 2000s, customer acquisition became relatively easy, particularly with mobile apps, the vast majority of which were offered for free download. While advertising has traditionally been the revenue model for B2C software companies, many of them have been opposed to ads for reasons related to product design and consumer preferences, and as a result, have generated little to no revenues while amassing large volumes of users. And yet, despite the absolute lack of revenues or any tangible assets (eg, product inventories), some of these companies have exited in valuations of up to billions of US dollars.

The key to understanding some of these valuation anomalies, which also helps to frame the data monetisation models that will be presented next, is by viewing users as bundles of data. Take, for example, two of the most prominent data-centric transactions driven by users – the 2012 acquisition of Instagram for $1 billion (with approximately 30 million reported monthly active users) and the 2014 acquisition of WhatsApp for $19 billion (with roughly 450 million reported monthly active users), both pre-revenue start-ups acquired by Facebook.

While there arguably may have been some value in the technology category of these two companies, these were both mobile applications operating on a fairly standard technology platform and it is unlikely that this was the basis for billions of dollars in valuations. The value was really embedded in the users, which are proxies for data and represent future monetisation options. Indeed, Facebook went on to realise significant returns on these users.

The enterprise as a steward of data

While users may as well be priced as valuable data bundles, one of the questions hampering the monetisation of user data – and other types of data collected by the enterprise – has been: who owns the data? This question is particularly challenging in IoT environments. Take, for example, a smart home device such as the Google Nest thermostat. The device is installed in private homes, collects information on ambient temperatures and user heating preferences and translates this into heating and cooling controls inside the home via the heating, ventilation and air conditioning system. Since there are multiple parties involved in the process, access to the data can be controlled by the following:

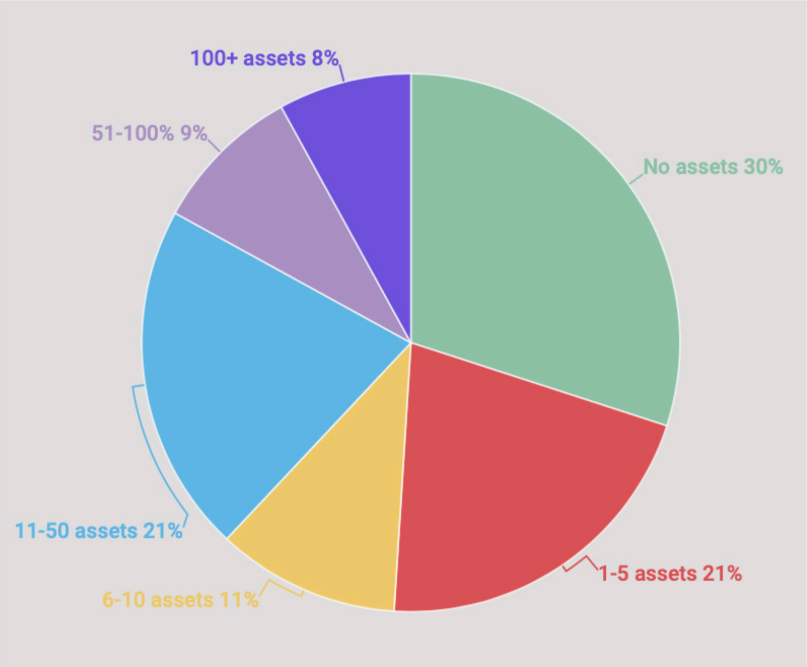
- the end user – the owner of the home where the thermostat is installed, who allows the collection of data required for the operation of the system;
- the hardware maker – Google, which makes the thermostat, and who stores all data in the Cloud to then be utilised by AI algorithms to control and improve energy consumption related to heating and cooling;
- the energy utility – provides the physical infrastructure for heating and cooling through gas and/or electricity and collects key data related to actual energy consumption; and/or
- the solar company – in case of a solar home, there will also be the solar company (eg, SunRun, which in most cases leases the system to the homeowner) and collects data related to energy generation through the solar panels.

The ambiguity surrounding data access and ownership associated with IoT ecosystems and other similar networks accentuates the role of the enterprise as a steward of data, a concept highlighted by IDC in its Data Age 2025 study. With the transition to cloud hosting and data management, more and more consumer data is collected and kept by enterprises with which they do business. The responsibility to maintain and manage all of this consumer and business data supports the growth in cloud hosting through data centres. As a result, the role of the enterprise as a data steward continues to grow.

Safeguarding consumer data touches on issues of security and privacy that often need to be regulated on a national level. In the healthcare field, the Health Insurance Portability and Accountability Act (1996) (HIPPA) is an example of a US federal law that requires the creation of national standards to protect sensitive patient information from being disclosed without the patient's knowledge or consent. Further, provisions in the act mandate the adoption of federal privacy protections for individually identifiable health information. The HIPPA rules apply to entities such as health plans and health care providers dealing with patient data.

More recent initiatives related to corporate data stewardship include the enactment of the General Data Protection Regulation (GDPR), the European Union's landmark data privacy and security law, which came into effect in May 2018. GDPR applies to organisations that process personal data of EU citizens or residents as well as organisations that offer goods and services to EU citizens or residents. Hefty fines are imposed on violators according to a tiered scale based on the severity of the violation. One of the most publicised provisions of GDPR relates to people's right to erasure, known as the right to be forgotten, which grants individuals the right to request that organisations delete their personal data. In the United States, the California Consumer Privacy Act (2018), which came into force on 1 January 2020, gives consumers more control over the personal information that businesses collect about them and secures new privacy rights for California residents.

FIGURE 4. US unicorn IP portfolio breakdown



Source: Foresight Valuation Group

TABLE 1. Worldwide security spending by segment, 2017 to 2019 (millions of dollars)

Market segment	2017	2018	2019
Application security	2,434	2,742	3,003
Cloud security	185	304	459
Data security	2,563	3,063	3,524
Identity access management	8,823	9,768	10,578
Infrastructure protection	12,583	14,106	15,337
Integrated risk management	3,949	4,347	4,712
Network security equipment	10,911	12,427	13,321
Other information security software	1,832	2,079	2,285
Security services	52,315	58,920	64,237
Consumer security software	5,948	6,395	6,661
Total	101,544	114,152	124,116

Source: Gartner (August 2018)

According to Gartner estimates, worldwide cybersecurity spending from 2017 to 2019 exceeded (or is expected to exceed) $100 billion annually (see Table 1). One Gartner study (2018) further projects that the cybersecurity market size will increase to $270 billion by 2026. Much of the growth in this spending emerges from the high economic cost of data breaches – a risk that is only expected to intensify post-covid-19, due to the increase in remote workforces. According to a 2019 IBM survey, the average cost of a data breach in the United States has more than doubled, from $3.54 million in 2006 to $8.19 million in 2019.

Data monetisation – the leading business models

Against this backdrop of exponentially increasing volumes of data being collected and processed on the one hand and strict data privacy regulations and mounting security threats on the other, data monetisation remains largely limited in scope. 'The State of Dark Data', a survey of 1,300 IT and business leaders conducted by data management platform, Splunk, revealed that 55% of the surveyed organisations' data is 'dark', which is defined as "untapped and, often, completely unknown". Yet, the vast majority of survey participants agreed that data is "extremely valuable for success".

While data is an intangible asset, its value – unlike patent value – is not driven by exclusionary rights. Moreover, unlike patent monetisation where there are usually two parties involved (the IP holder and the IP user), data monetisation is often carried out via three-way models, where data is collected in the course of providing services and is then monetised via a third party.

As a result, the business models surrounding data assets cannot be entirely based on licensing or enforcement as the unique nature of data assets, particularly when leveraged through AI algorithms, can bring value to the enterprise in several novel ways. Generally speaking, there is a distinction to be made between passive and active monetisation, which will both be covered in connection with data assets, but the focus will be on active monetisation, where much activity has taken place in recent years.

Passive data monetisation

Passive monetisation of intellectual property can be achieved through the collateralisation of IP assets, most commonly patent portfolios, against debt financing (ie, loans). IP collateralisation activity in general is far from being in the mainstream, as most commercial banks are conservative financial institutions, bound by traditional loan ratios and other metrics that do not factor in the value of IP assets. Since these assets are not reported on balance sheets, they are missing from the accounting definition of 'book value'. There is also a gap in reporting IP transactions, which creates a valuation comparables void and thus makes it difficult to value intellectual property as collateral.

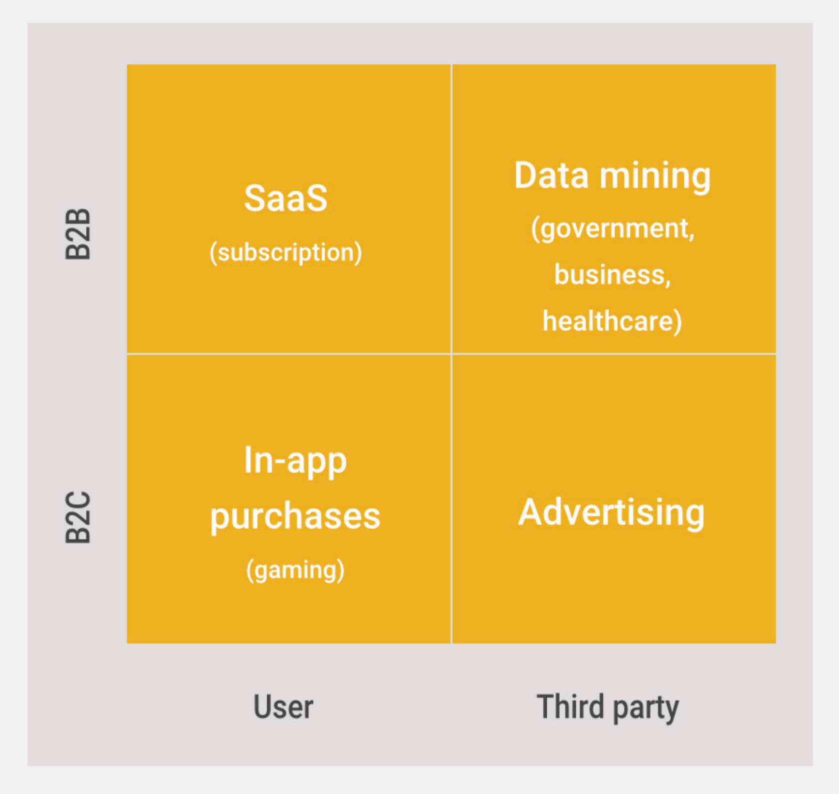
When it comes to data assets, these cannot easily be leveraged for funding or even collateralised in ways that patents can, due to both the lack of the exclusionary IP protection that patents offer and the difficulty associated with data valuation. Several data collateralisation mechanisms have emerged in recent years. One that stands out is Leeward Capital Management with its Sale-to-Service (S2S) offering. This is similar to a sale-leaseback but rather than buying tangible assets, such as equipment or real estate, Leeward acquires a company's systems, processes and data, which reside on a server in a data centre or in the Cloud.

Active data monetisation

The matrix in Figure 5 presents a novel framework for mapping out the various business models associated with active data monetisation. This is a dynamic model and is updated frequently based on experience gained through client projects and observations in the market.

This framework is based on the type of customer – business to business (B2B) or business to consumer (B2C); and the monetised party (ie, who is paying for the product/service) – user or third party. Identifying the monetised party is critical, due to the proliferation of three-way monetisation schemes. The underpinnings of each model will be discussed, with examples of companies or sectors in the market that have successfully implemented each business model.

FIGURE 5. Data monetisation – business models



	User	Third party
B2B	SaaS (subscription)	Data mining (government, business, healthcare)
B2C	In-app purchases (gaming)	Advertising

Source: Foresight Valuation Group

SaaS and advertising models – the status quo

The SaaS business model and the advertising business model are two of the most commonly applied data monetisation schemes in the market today.

Software-as-a-Service (SaaS)

The SaaS model is a subscription model common in B2B situations where the user (a business) is paying for access to software or data. This model is based on an access fee and is the most comparable to patent licensing of the four active data monetisation models presented. One example of SaaS access to data is the digital subscription service LexisNexis, which provides online access to case law and other legal information via a monthly subscription. Unlike patent licensing, the data accessed by LexisNexis subscribers is not proprietary – it is aggregated through public sources (some of it may be copyrighted to publishers who used to aggregate it in books, prior to the availability of digital access).

The subscription fee for this type of service can be viewed as a convenience fee – legal information is voluminous and paper access is becoming impractical. According to its most recent annual report, the LexisNexis legal and news database contains 119 billion documents and records, which include 250 million court dockets and documents. The subscription business model has seen a significant shift over the past 10 years as the legal services market has shifted from print to rely on online access.

Monetising data via a SaaS model is one of the most straightforward business models and can fit almost every industry where access to large amounts of data is necessary, including sectors such as agriculture, transportation and biotech. Its advantages are in ease of delivery and access, the recurring nature of revenues (eg, metrics, such as monthly recurring revenues, are frequently tracked), the ease of upselling additional products and offerings to existing customers and the convenience of real-time updates (eg, in the case of LexisNexis, this replaces the need to buy new hardcover editions every year). The main disadvantage of the SaaS business model is the risk of customer churn, which can be measured by the percentage of existing customers leaving every month. Since customer acquisition cost is spent upfront in sales and marketing and revenues are realised in small monthly increments over time, high turnover is not a desirable outcome since it reduces the lifetime value of a customer – a key success metric in SaaS.

Advertising

The advertising model has been the most common monetisation scheme since the dawn of the Internet and is particularly common with B2C applications. This model is based on a three-way monetisation system – consumers access online applications (web or mobile) for free, their data is aggregated and made accessible to third-party advertisers, who essentially cover the free service via advertising spending on the app or website. The monetised party is not the user, but a third party (advertising brand). In order for this process to work efficiently, there are sophisticated advertising networks serving the ads and other technology infrastructure that facilitates the matching of customers and messaging. Both Google and Facebook generate most of their revenues from advertising, based on this general three-way model: Google generated more than 83% of its $161.9 billion revenues in 2019 from advertising across its various platforms; likewise, Facebook generated $69.7 billion from advertising in 2019, more than 98% of its total revenues for the year.

Advertising is one of the most ubiquitous data monetisation models as almost every free B2C app has some component of advertising revenues supporting its operations. Its advantages are in its simplicity and self-propelling nature – the ability to offer free access attracts more users who, in turn, provide more and more data, which then attracts more advertisers. While classified as an active monetisation scheme, this type of model runs on autopilot as long as there is significant traffic to a site or app. The key disadvantage of this model is that it constantly tests the boundaries of consumer privacy and data stewardship. With the advancement of sensors on mobile devices and the ability to capture sensitive data such as biometric information, privacy concerns intensify and become a target of government intervention, which could impede monetisation going forward.

While businesses traditionally pay for SaaS services and consumer data is traditionally monetised through advertising, we are seeing a convergence through some hybrid business models where consumer SaaS is appearing. This type of hybrid is particularly common in wearables – a segment of the IoT market. Wearable devices (eg, smart glasses, smart watches or any other connected device worn on the body that can take vital measurements) collect health data that consumers may be interested in paying access for. The business model for some of these usually includes a free tier of basic access to data and a paying tier (a model known as 'freemium') of access to things such as data history over time, data analytics and nutritional recommendations. There may also be advertising on top of this, so these apps also include the three-way monetisation that is common for B2C services (albeit at the odds of running into regulatory challenges, which are much higher with health data).

The next frontier

The in-app purchases model and the data mining model are the more innovative models on the active data monetisation matrix and both are still shaping up and evolving in the marketplace. These represent the future of data monetisation.

In-app purchases

The in-app purchases model is one of the few instances where data monetisation is taking place at the consumer level. It is a two-way model, involving the consumer as the paying party. While consumers do not like to pay to download apps (as Apple CEO Tim Cook recently testified, 84% of apps on the Apple App Store are free), gaming apps are an exception to the rule. In some games, players can pay for digital currency that allows them to buy accessories in the game, also known as 'cosmetics' (modifiers that change the way that certain objects look in the game). Gaming is an example of the types of model that include the monetisation of digital assets, an extension of data into other digital commodities, which form a new class of digital intangibles. These models involve the use of digital assets either as payment mechanisms, such as tokens (common currencies in blockchain decentralised networks) or as the goods being acquired in virtual environments. In-app purchases made headlines when Epic Games, publisher of the hugely popular game Fortnite (which has allegedly been downloaded on the Apple App Store nearly 130 million times) announced in August 2020 that a new direct payment option for players is available to purchase the currency used in the game outside of the iOS App Store or Google Play. This direct payment option cut Apple and Google from their revenue share (30% of all app-related revenues) as the transaction would not go through their respective platforms. In response, Apple and Google pulled the app from their app stores for violations of their terms of service and Epic subsequently filed suit against both companies alleging antitrust violations.

The legal battle surrounding Fortnite shows the pros and cons of this data monetisation model. On the pro side, it has appealing economics as it provides revenues from sales of virtual goods with no cost to create or deliver. Consumers do not pay for the app, but they pay for the virtual goods, so it taps into consumer behaviour in a very powerful way. However, the thorn in the otherwise appealing profit margin opportunity is the high cost of the carrying platform, as embodied in the 30% charged by Apple, which gave rise to Epic's legal battle. It will be interesting to watch how this situation unfolds as the market will need to find an equilibrium that works for both sides; the role of the platforms is critical in distributing the game, but at the same time, 30% of revenues may be a bit steep for the game publishers.

Data mining

Finally, the most ambitious model on this map is the data mining model, which represents the holy grail of data monetisation on a corporate level. This is where the market has not quite figured out all the possibilities, as issues of data ownership, security and privacy are major hurdles to fully realise the potential of data mining. This is a multi-party monetisation model, involving large-scale data collected across industries, devices and physical environments.

The pioneers in data mining are governments and healthcare systems, who have access to data at a large scale and use predictive analytics and other tools to drive public health policy (eg, in the covid-19 pandemic) or for national security purposes. The scale of data collection and analytics involved here are often beyond the capabilities of most government agencies or corporations, so what has emerged in the market are intermediary platforms that process the data and share the results with customers under various arrangements.

One data mining platform that stands out is Palantir, which recently filed for an initial public offering, providing a rare glimpse into its highly secretive operations. According to Palantir's prospectus, its software platforms are used by many of the world's most vital institutions, from defence and intelligence agencies to companies in the healthcare, energy and manufacturing sectors. Palantir offers two software platforms, Palantir Gotham and Palantir Foundry. The former was built for commercial institutions to create a central operating system for their data, while the latter was constructed for analysts at defence and intelligence agencies who were "hunting for needles in not one, but in thousands of haystacks". In the first half of 2020, Palantir's platforms were used by 125 customers, including the US Army. The company's prospectus provides just a hint of the benefits derived by their customers, stating that their pricing is based primarily on the expected value that their platforms produce for their customers. Other companies that provide similar services include Tableau, Cloudera, Teradata and Qlik.

Action plan

A

In order to be data-monetisation ready, organisations should make sure that they have all measures in place to allow them to leverage their valuable digital intangibles for maximum return, while minimising security and regulatory risks:

- Keep your corporate data well protected in the Cloud or on-site, to prevent data breaches.
- Review your trade secret protocols to make sure that access to your data is well covered by legal contracts and other required measures.
- Ensure compliance with jurisdictional data privacy laws, such as GDPR.
- If you are in the B2B space, you should already be engaging in some form of data monetisation: explore ways to utilise platforms in the market to get more insights from your data incorporated in your business decisions, or even try to package some of the data that you may be able to extend via a SaaS model to other companies/industries.
- If you are in the B2C space, do not be deterred from engaging in data monetisation activities: advertising is not the only way to monetise consumer data, there may be creative models, like the creation of virtual goods, to engage in additional monetisation from your users.

Comment

The role of the organisation as a 'steward of data' should not be misconstrued as an injunction on data monetisation. On the contrary, data has the potential to enhance corporate value in significant ways, and data assets should be viewed as an integral part of the modern IP portfolio (as digital intangibles). Using the analogy of data as the new fuel, the engines of data analytics are already revving up, all that is needed is to properly address roadblocks such as privacy and security. Every tech company that wants to jump on the data monetisation train should be very familiar with the types of models and platforms allowing it to leverage its data assets in either a two-way or a three-way model and eventually embark on the full benefits of data mining when the time is right.

Efrat Kasznik is the president of Foresight Valuation Group, LLC