

How to Extract Value from Data?

*By Analyst Oliver Hinsch
C WorldWide Asset Management*

Data has shaped human decisions long before computers, the internet, and AI emerged. The first evidence of data collection and analysis dates back to 18000 BC. In modern-day Congo, the Palaeolithic tribespeople marked lines into bones. The bones, known as Ishango bones, were used to monitor trading activity and forecast the longevity of food supplies. As societies evolved, libraries became the first attempts at mass data storage. Fast forward to the 21st century and the data explosion, where AI and LLM are expected to be catalysts for extracting even more value out of data. No doubt, data will continue to grow rapidly, but the big question is which companies will be able to extract long-term value from the data?

In our analysis, we differentiate between pure data companies and companies exploiting data to complement their existing business. Among the pure data companies, we distinguish between "Standard-bearers" and "Data librarians".

We conclude that while some companies are favourably positioned to capture long-term value from a data-driven business model, for the majority, data will be a prerequisite to remain competitive and not necessarily a long-term driver of shareholder value. Across our strategies, we currently have portfolio exposure to five data companies: RELX, Verisk, S&P Global, Visa and Alphabet.

Data is a growing industry

The amount of data libraries could store was constrained by physical limitations. Scarce capacity made data expensive to store, while its physical nature made it difficult to utilise. Only with the invention of magnetic tape, semiconductors, and computers did it become possible to store data digitally. By 1996, digital storage had become more cost-effective than paper. The amount of data produced grew exponentially with the gradual adoption of the internet, sensors, and digital tools. Data became a byproduct of everything, generated from digital and physical activity in transport, production facilities, and weather patterns. As show in Figure 1, digital data has grown by 60% annually since 1999 and 48% since 2006. With lower processing costs, companies became better at managing and using data, increasing demand for storage and tools for managing data. Industry experts expect the amount of data produced to grow 10x between 2020 and 2030.

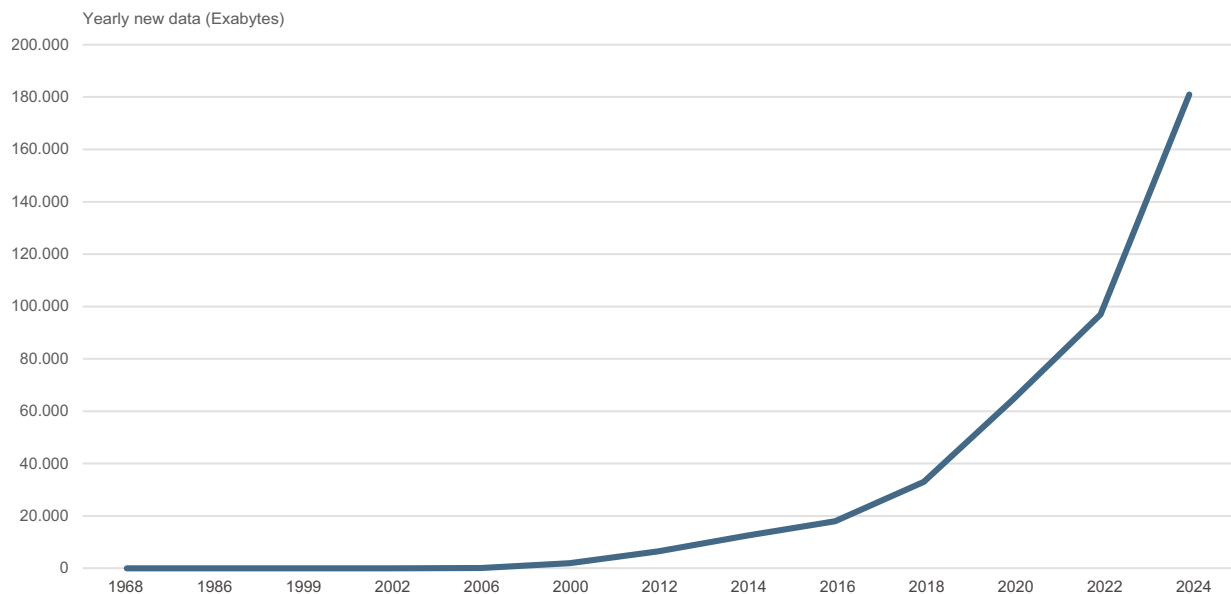
Evolutions in data have made certain business models irrelevant, enabling new ones to emerge and others to

evolve. While most companies use data, only a few have built a profitable business model around monetising the data itself or the insights it provides. Platform companies like Alphabet, Amazon, and Meta have successfully capitalised data through targeted advertisement. Other businesses have been accumulating large amounts of proprietary and trusted data for decades. As data became digital, the latter companies could extract more knowledge from the same data. Furthermore, some traditional industrial companies have pivoted towards extracting data and knowledge from their large installed base, installing sensors, and building intelligent solutions, a theme we call “[The intelligent tangible world](#)”.

Data is abundant, but not all can gain a sustainable competitive advantage

Data collection and interpretation improve the decisions and make resource allocation more efficient. This trend has only continued with stronger computing power. Although data is important, it is not the new oil. Only a few

Figure 1
An explosion of new digital data



Source: Michael Lesk (1998), Peter Lyman (2000), The Economist (2010), Wonder(2012), and Statista (2024).

Note: Because there is no unambiguous way to measure the size of digital information, definitions vary throughout time and may be either inconclusive or estimates. 1 Exabyte = 1,000,000,000,000 MB or +222,000,000 full-length movies.

companies can process data at a scale or with a brand that increases entry barriers and cements permanent competitive advantages. Like oil, data will likely fuel the digital economy. That doesn't mean every company will be an oil company. Rather, data is more like sand. It is available almost everywhere, but only a few companies can make it into silica, the raw material of silicon used in semiconductors.

AI and LLMs do not change this. Businesses should use data to improve efficiencies and outcomes. While more data is generally better, and LLMs boost the capacity by which data can be extracted and analysed, the degree of

obsession and expectations are likely to disappoint. For most companies, advantages gained through data will not provide comparative advantages. They will provide efficiencies, but competitors will likely mimic those. Like the effects data analytics had on basketball (see separate box), once every team optimised their playing style, all teams returned to a similar baseline. This is also true for the effect data will have on most companies. In certain cases, however, one company is the sole or dominant provider of relevant industry data. These companies can earn a sweet spot position and earn a toll for usage. We call these data companies.

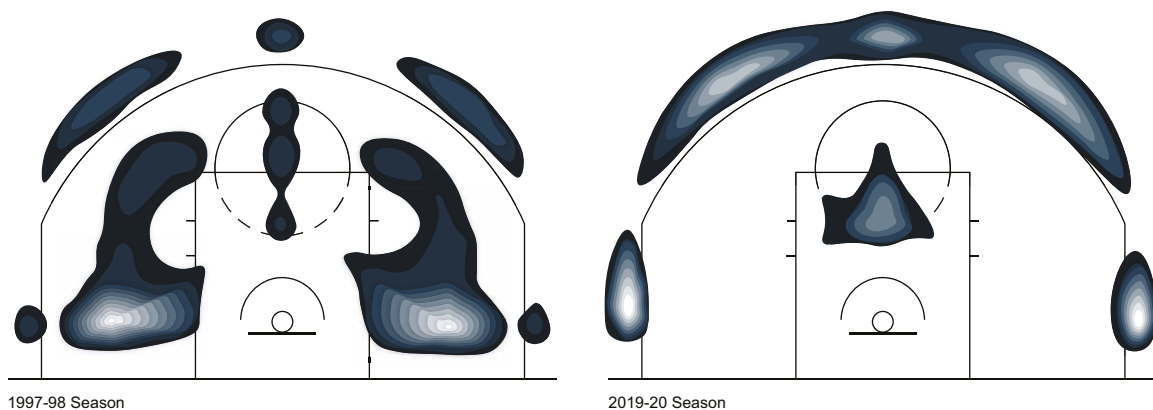
“Data is like sand. It is available almost everywhere, but only a few companies can make it into silica, the raw material of silicon used in semiconductors.”

Case Study: Basketball

Data can underpin knowledge and change behaviour

Data and data analytics convey knowledge that can change behaviour. These effects can be seen in everything from how businesses inform pricing decisions to how basketball is played. Kirk Goldsberry famously observed the drastic changes in how basketball was played between the early 2000s and the early 2020s. Spatial tracking and data analytics enabled detailed analysis of, for example, the rationale for attempting a shot behind the three-point line. The simple conclusion: the value of attempting to score three rather than two points outweighed the increased difficulty of moving a few meters out of the mid-range zone. Data analytics changed basketball. Today, whole teams and offensive attacks are structured around the three-point shot.

Most common jump shot locations in the NBA



Data companies can have attractive unit economics

Data becomes valuable when the producer has unique access to valuable data or when consumers agree that data from one company is the single source of truth (i.e. it becomes the industry standard). Once established, these businesses often become monopolies or duopolies in their business area. Dominant data companies typically have broad distribution and manage a large amount of data, underpinning better offerings and stronger unit economics. This network effect cements high barriers to entry and winner-takes-most markets. As companies become dependent upon data, data companies become deeply entrenched in the workflow of customers through data analytics. This underpins the ability of these companies to effectively cross and upsell new products to customers, establishing attractive reinvestment opportunities.

Once data is gathered, cleaned, and optimised, it can be sold repeatedly without incurring additional costs. Additional sales generate nearly 100% incremental margins. While operating margins are initially low, significant op-

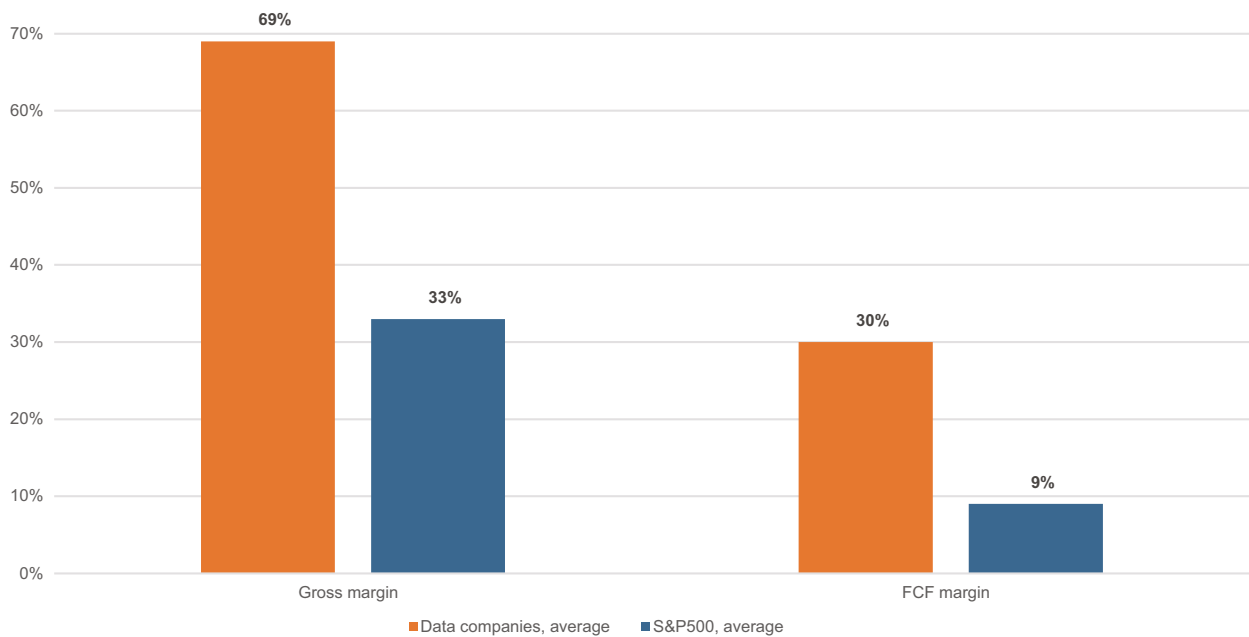
erating leverage enables meaningful margin expansion as these companies scale. We monitor a select group of data companies. These have average gross margins of 69% and FCF margins of 30%, compared to 33% and 9% for the S&P 500, as illustrated in Figure 2 below.

Standard bearers and Data librarians

Two types of data companies embody these characteristics – Standard bearers and Data librarians. Standard bearers have a unique brand that makes their data more valuable because it has become accepted as an industry standard. This is like the meter, the standard through which distance is measured and communicated, or the TEU, which is the size of shipping containers. Data librarians have unique access or processes around data collection. Their data is either impossible or very cumbersome to replicate.

Standard bearers utilise various data sources to create benchmarks upon which whole ecosystems rely. For example, MSCI is the leading global provider of index data and analytics tools for the asset management industry.

Figure 2
Gross margin and FCF margin, data companies



Source: Bloomberg (2024) and C WorldWide, November 2024.

“Two types of data companies stand out: Standard bearers and Data librarians. Standard bearers, whose data gain value by becoming the industry standard, and Data librarians, who have exclusive access or processes that make their data hard to replicate.”

MSCI is synonymous with measuring and describing the performance of global equity markets. While asset managers pay for MSCI data, asset owners, such as pension funds and endowments, require external managers to report against a benchmark they know and trust. It is, therefore, not the data itself that is valuable; it is the MSCI brand. The underlying data can be replicated, but the deep integration into contracts and workflows of asset owners, decades of consistency and brand-building, and trust in their data specifically are challenging to replicate.

Two other Standard bearers are Moody’s (credit ratings) and S&P Global. S&P utilises credit information, stock prices, and other publicly available data, turning them into benchmarks. S&P is the standard bearer of how corporate credit quality is communicated (ratings), how the performance of the US equity market is tracked (S&P500), how commodity contracts are settled (Platts), and how vehicles are registered (Carfax). S&P has bolstered these standards with complementary analytics.

Data librarians have unique access to data. One such example is Verisk Analytics. In 1971, new regulations in the US required insurance companies to share data with state regulators in a standardised and accurate format. Rather than each building their own, 280 insurance companies consolidated their data aggregation and cleansing in a not-for-profit called ISO. Due to its initial success, insurers contributed claims and loss data to the consortium. This data was standardised and shared with all members, forming the basis of fraud detection, pricing decisions, and risk assessments. Verisk Analytics, established as the parent company of ISO in 2008, remained non-profit until 1997 and became a public company in 2009. Starting as a data company, it developed analytical products utilising the shared data. Today, Verisk Analytics operates in mo-



nopolies and duopolies within the US property insurance industry, banking, and energy markets, serving data and the analytics around it.

RELX is another example of a Data librarian. RELX has unique and dominant data assets within various industries. RELX traces its origins to Elsevier, a news magazine founded in 1880. In five decades, Elsevier built and acquired industry-specific magazines and journals, ultimately becoming the largest B2B publisher in Europe. RELX operates leading positions in oligopolistic data and business analytics markets within the legal industry, US auto insurance, banking, security service, aviation, medical research, academic publishing, and chemical pricing.

The professional readers of the journals often contribute the proprietary data. In 1971, Elsevier became the first company to store journal information in a computerised database. This transitioned Elsevier from a media company to an analytics company, charging customers for access to data and analytics instead of journals. Clients have few alternatives and rely on this data for critical decisions and operations; therefore, RELX can enjoy potential pricing power around the data.

In Table 1 we have listed some of the most prominent companies that can be characterised as Standard bearers and Data librarians.

Data companies also have risks

Most data companies have existed for decades, even centuries, and built strong barriers around their right to win. Despite the seemingly insurmountable structural factors and dynamics underpinning their strength, history has shown that even these can fail. As Benedict Evens, an independent technology analyst and former partner at Andreessen Horowitz, puts it, a company’s structural competitive advantage can either be ordered by a king to be knocked down or what it protects may become irrelevant (Benedict Evens: “How to lose a Monopoly”, 2020).

The king, often a regulator, may have granted a company its strong position. If the company acts irresponsibly or anti-competitively, the king can choose to break down the barriers of protection. Dun & Bradstreet (D&B) was given a monopoly in 1996. In 1996, the federal government required all companies working with the government or filing certain documents to have a DUNS number. The DUNS number, operated by D&B, is an individual identification number for businesses, linking ownership and trade credit data to specific companies. This established DUNS as a standard. Many businesses adopted the DUNS

“Despite the seemingly insurmountable structural factors and dynamics underpinning companies with decades of experience and strength, history has shown that even they can fail.”

Table 1
Examples of prominent data companies

Standard bearers	Data librarians
FICO	Experian
MSCI	Transunion
S&P Global	Equifax
Moody's	RELX
CME	Verisk
ICE	LSEG
Gartner	Deutsche Börse
Nielsen Holdings	Costar
IQVIA	Factset
	Morningstar
	Visa
	Mastercard
	Meta
	Alphabet

Source: C WorldWide, November 2024.

number to determine if companies were eligible for loans and compliant business partners. This put D&B in a sweet spot. While it remained a standard, D&B did not transition into analytics. In 2022, the government moved from the DUNS number to an open-source code, breaking down the most entrenched barrier.

Another risk is customers destroying the competitive advantages. RELX and Verisk receive data from customers. This data is cleaned, refined, and sold to the same clients. Contributory data models are very powerful but hard to achieve. While these are often natural monopolies, they serve as a common good for the industry. If they exploit high pricing power by appearing as toll-takers, customers can create a new data consortium. In the late 1990s, RELX increased prices significantly within its academic journals business. In this business, RELX receives the academic work of researchers and publishes it in journals. Price increases resulted in a backlash from both contributors and clients. This led to the emergence of alternative forms of distribution – most notably open access. While RELX remains the leader today, it recognised that unutilised pricing power may be even more powerful for contributory data models than realised pricing power. Lower prices incentivise the adoption of new products, enabling more

data and thereby generating more customer value, further entrenching RELX into the workflow of customers.

The source around which a company has built a competitive advantage may also run dry. Competitive advantages can become irrelevant, not because they are diminished, but because the lay of the land has changed, rendering the very thing it protects immaterial. Nielsen Holdings was an aspirational data company. Founded in 1923, it coined the very concept of market share and became ubiquitous with TV ratings, retail share, and market data. Nielsen had access to both valuable proprietary data and was the standard on which USD 80 bn+ of television advertising spending was based. Nielsen had a network of 100,000 Americans carrying a small device when watching TV. The company also embedded a digital watermark in the audio of television programs, enabling the device to recognise when, how much, and which television programs were watched. It became the standard bearer, fuelling the TV advertising market. However, the internet changed the lay of the land. Consumers migrated towards streaming and mobile phones. Advertising moved to Facebook, YouTube, and Google. The media landscape became fragmented. Although Nielsen remains the standard bearer for TV ratings, the medium itself became less relevant.

Finally, the sensitive nature of information makes data companies a target for hackers. This reinforces the political emphasis on data sovereignty and privacy. Data sovereignty may ultimately limit data companies' growth opportunities and business development. Data privacy

regulations like the GDPR increase the implicit storage costs of holding and transacting with data, and with this follows the risk of increased political scrutiny.

Data is an interesting fishing pond

Data is, without doubt, growing and becoming more important to all businesses. AI and other tools will further enable more valuable insights to be extracted from data, effectively increasing productivity for most companies. For example, leading industrial companies are utilising sensors to extract data from their machines and selling such data to their customers. As discussed in other Insights, this solidifies their market positions with intangible intelligence around their products.

In this paper, we looked at companies that have made data the backbone of their business model. We make the distinction between standard bearers and data librarians. Both types of data companies often develop into natural monopolies with deeply entrenched rights to win, very attractive unit economics, high margins and returns on capital with strong underlying growth drivers and attractive reinvestment opportunities.

As long-term stock pickers, we see this area as an attractive fishing pond. However, selectivity is required as the area is not without risks. We currently have exposure across our strategies in five of the companies mentioned above: RELX, Verisk, S&P Global, Visa and Alphabet.



Disclaimers

EU: This is marketing material. This publication is prepared by C WorldWide Asset Management Fondsmæglerselskab A/S. It is provided for information purposes only and does not constitute, and shall not be considered as, an offer, solicitation or invitation to engage in investment operations, as investment advice or as investment research. The publication has thus not been prepared in accordance with legal requirements designed to promote the independence of investment research, and it is not subject to any prohibition on dealing ahead of the dissemination of investment research. Opinions expressed are current opinions only as of the date of the publication. The publication has been prepared from sources C WorldWide Asset Management Fondsmæglerselskab A/S believes to be reliable and all reasonable precautions have been taken to ensure the correctness and accuracy of the information. However, the correctness and accuracy is not guaranteed and C WorldWide Asset Management Fondsmæglerselskab A/S accepts no liability for any errors or omissions. The publication may not be reproduced or distributed, in whole or in part, without the prior written consent of C WorldWide Asset Management Fondsmæglerselskab A/S. Past performance does not indicate future performance.

US: This is marketing material. This publication has been prepared by C WorldWide Asset Management Fondsmæglerselskab A/S (CWW AM). CWW AM is a registered Danish investment firm located at Dampfaergevej 26, DK-2100 Copenhagen, Denmark. CWW AM's Danish company registration no is 78420510. CWW AM is registered with the SEC as an investment adviser with CRD no 173234. This publication is provided for information purposes only and does not constitute, and shall not be considered as, an offer, solicitation or invitation to engage in investment operations, as investment advice or as investment research. The publication has thus not been prepared in accordance with legal requirements designed to promote the independence of investment research, and it is not subject to any prohibition on dealing ahead of the dissemination of investment research. Opinions expressed are current opinions only as of the date of the publication. The publication has been prepared from sources CWW AM believes to be reliable and all reasonable precautions have been taken to ensure the correctness and accuracy of the information. However, the correctness and accuracy is not guaranteed and CWW AM accepts no liability for any errors or omissions. The publication may not be reproduced or distributed, in whole or in part, without the prior written consent of CWW AM. All figures are based on past performance. Past performance does not indicate future performance. The return may increase or decrease as a result of currency fluctuations.

UK: This document has been prepared by C WorldWide Asset Management Fondsmæglerselskab A/S (CWW AM). CWW AM is a focused asset manager registered as an investment firm with the Danish FSA. CWW AM is located at Dampfaergevej 26, DK-2100 Copenhagen, Denmark, CVR registration number 7842 0510. This document is directed at persons having professional experience of participating in unregulated schemes (investment professionals) and high net worth companies (as defined under art. 14 and 22 of the Financial Services and Markets Act 2000 (Promotion of Collective Investment Schemes) (Exemptions) Order 2001). The document and any investment or investment activity to which it relates is available only to such persons and will be engaged in only with such persons. Any other person should not rely or act on the statements made in this document. The content of this presentation is confidential, and redistribution or reproduction is prohibited. The presentation is provided for information purposes only and does not constitute, and shall not be considered as, an offer, solicitation or invitation to engage in investment operations as investment advice or as investment research. CWW AM is not responsible for the suitability of the information in the presentation. Opinions expressed are current opinions only as of the date of the presentation. The presentation has been prepared from resources CWW AM believes to be reliable and all reasonable precautions have been taken to ensure the correctness and accuracy of the information. However, the correctness and accuracy are not guaranteed and CWW AM accepts no liability for any errors or omissions. This presentation contains information on investments which does not constitute independent research. Accordingly, it is not subject to the protections afforded to independent research and CWW AM, and its employees may have dealt in the investments concerned. It is emphasized that past performance is no reliable indicator of future performance and that the return on investments may vary as a result of currency fluctuations.

C WORLDWIDE ASSET MANAGEMENT FONDSMAEGLERSELSKAB A/S

Dampfaergevej 26 · DK-2100 Copenhagen

Tel: +45 35 46 35 00 · Fax: +45 35 46 36 00 · VAT 78 42 05 10 · cworldwide.com

Q4 2024