Opinion **Big Data**

China is far ahead on drilling for data

Allegations of information theft in Houston point to the importance of 'the new oil'

LEO LEWIS

© Ingram Pinn/Financial Times

Leo Lewis 10 HOURS AGO

Since it was first proposed in 2006, the contention that "data is the new oil" has gradually charmed its way into mainstream discourse as a neat and comprehensible explanation of something messy and incomprehensible.

The equivalence, however stretched, is alluring. The idea that the modern world has conferred stupendous commercial value on data feels validated by daily experience: spookily targeted social media adverts, shoemakers tracking your morning run, algorithms in the workplace that learn their trade from people-generated data sets, US congressional antitrust hearings on tech giants with stratospheric valuations.

The analogy's geopolitical accuracy also appears to be confirmed by threatened or real bans of TikTok and Huawei. It is easy to talk data up as the indispensable lifeblood of modern economies. As with oil, we can see its protection becoming a matter of national priority, long-term strategies, and conceivably, conflict.

The now pressing questions for business — across all industries and around the world — are how literally to take the analogy, whether we need more assiduous protection of data and whether Chinese companies, especially those backed by the state, made a devastatingly correct call on this ages ago.

The issue, argue people closely involved in detecting Chinese data theft from western companies, reached a critical, but imperfectly understood, turning point last month. That's when the US ordered China to close its consulate in Houston. Many have suggested the move signifies an intractable ideological rift and the start of a new cold war. Less obviously, it also suggests that data, primarily that produced and gathered in vast quantities by private companies, may have finally joined oil as a *casus belli*.

The US State Department's explanation for the closure was that it had been undertaken to protect American intellectual property and Americans' private information. That justification dovetails with existing claims that Chinese hackers are stealing US intellectual property — from pharmaceuticals, tech and military hardware to entire industrial processes — that may run into the hundreds of billions of dollars every year.

But people whose work contributed significantly to the Houston decision tell me that the specifics on what was happening there support those analysts who argue that the consulate closure was not just political but also a practical response to commercially-driven larceny.

Much of what was being spirited out of the US via China's Houston consulate, these investigators allege, was raw data generated by the great army of energy and agricultural commodity firms clustered in America's fourth city. Granular information was being siphoned out of companies involved in investment, production, distribution and trading of energy and other commodities.

The information was wanted because it could provide Chinese companies with a hard commercial edge in the global energy and agricultural markets where China is the world's largest consumer. The project appears to have been a meticulous long-term effort, systematically targeting data both for its quantity and quality, the investigators say.

One former intelligence source, based in Houston and involved in cyber protection for energy companies, argues that there has been the significant imbalance between Chinese attitudes towards data and their western corporate counterparts. Some segments of the US economy do thrive on a sophisticated understanding of the value of data, led by Big Tech groups such as Amazon, Google, Facebook and Apple. And individuals are becoming aware of the monetary value attached to their personal data.

But in many industries, he says, the sense that information matters and must be protected remains stunted. This is particularly true for energy and commodities businesses, where the prevailing view is that oil remains both the old and new oil.

Attitudes may now change. But many people are reluctant to embrace the data-as-oil narrative because it is obviously flawed, says Alec Stapp, director of technology at the Progressive Policy Institute in Washington.

The list of fundamental differences between data and oil is considerable. Unlike data, oil can only be consumed once; it is fully fungible, its value is knowable before it is extracted and its production involves high marginal costs. There may also be excessive hype attached to the value of data: many data sets that could theoretically impart great value are — at least for now — unusable, or impossible to combine meaningfully with others.

There is a danger, however, in drawing comfort from this interpretation. While western companies allowed unprecedented volumes of data to escape, their Chinese corporate rivals have spotted perhaps the most important flaw in the data is oil theory. Information, unlike crude, costs very little to store in epic quantities. Chinese companies can afford, in that context, to play the long game and wait until the value of their existing data loot can be extracted and exploited. They may still be sitting on a gusher, even if the Houston consulate closure does mark the point at which western companies woke up to what they were losing and now rush to stem the flow.

leo.lewis@ft.com