Data Sovereignty and the VMware Sovereign Cloud Initiative
Table of Contents

Digital Sovereignty ................................................................. 3
Global Data Privacy Arena ....................................................... 5
Not All Data Is Created Equal .................................................. 7
Sovereign Cloud .................................................................. 8
VMware Sovereign Cloud Initiative ......................................... 8
Sovereign Clouds Augment Hyperscale Public Clouds .............. 10
Where to Next? ................................................................. 11
References ........................................................................ 11
Data sovereignty. Such a simple topic. Right? Drop my data just here in my local state, my local country, utilising a local availability zone or data centre for my hyperscaler-of-choice … I’m done, and all the locality concerns are addressed? Right? Well … there’s a lot more to it than that!

What may appear to be a relatively simple concept has significant considerations and ramifications for public and private organisations alike, as well as their citizens and consumers respectively. Across such public or private organisations, an initial question may be on the sensitivity of the data. If there is a degree of sensitivity, then that data needs to be appropriately protected. Now, how sensitive is the data? Well, I’m not sure! So, how do I protect it? What is the adequate protection of the data that I need to consider? As we continue to delve deeper into the question of the sensitivity of the data, what additional questions does the protection of my data … your data … our data … a public organisation’s citizen data … a private organisation’s consumer data … individual data … protecting all this data, what questions does it start to raise? Is this suddenly Pandora’s Box?

Well, it could be! The depth of the discussion around data sovereignty goes far beyond technology—it encompasses human rights, data privacy, national security, economic growth, the burgeoning data economy, national identity, the value of data, technological capability, artificial intelligence and machine learning … the list goes on.

Initially, let us briefly focus on the data privacy aspects of data sovereignty, then we shall touch on the relevancy of sovereign cloud as a capability that aids in addressing the data privacy concerns for sensitive or critical data and any artefacts derived from processing that data. Focusing on the data privacy aspects gives specific reference to the right to control the data, who has access to the data, where is the data located, who can modify or delete the data, and who has jurisdiction over that data.

We explore the definition of data sovereignty a little later in more detail when we specifically look at sovereign cloud, but for now let’s have a simple definition: the ability to maintain legal control and authority of the data (including any data flows, the subsequent processing of the data and any resulting data findings) within a jurisdictional boundary. Health data collected from Australian citizens during the COVID-pandemic from a mobile tracking app being maintained, accessed and controlled solely within Australia’s legal boundaries would be an example of data sovereignty. Further, consideration of the data sovereignty for the base health data collected by such a mobile application is just the start of the conversation or the tip of the iceberg; as this data is aggregated, processed and extrapolated, then these extended data sets have even more value (both economic and social) and the sovereignty of those extended data sets is a critical concern.

**Digital Sovereignty**

Beyond the consideration of data sovereignty, there is also the concept of digital sovereignty, and whilst we will not explore it to any great length here, it is important to note that there is significant global consideration of the aspiration to achieve digital sovereignty, well beyond “simply” the realisation of a nation’s or region’s data sovereignty. In this context, data sovereignty is a key subset of digital sovereignty, where a nation or jurisdictional region recognises the importance and the social value of having control of its “digital destiny”, realising and benefiting from end-to-end sovereign digital capabilities.
Data Sovereignty and the VMware Sovereign Cloud Initiative

An overarching aspiration of digital sovereignty represents the need to ensure a national (or regional), sovereign end-to-end digital capability and control. Such an aspiration is not only concerned with the sovereign protection of data, but going beyond such data sovereignty concerns and ensuring access to sovereign digital tools to develop applications, processes, manipulating the data and gaining insights from that data that benefit the nation and its citizens either directly for public organisations or indirectly through private organisations. Benefits extend to stronger national economic development and realising the significant social value that comes about from having control of the nation’s digital destiny.

Taking control of the digital destiny—aspiring to digital sovereignty—is about achieving digital autonomy with the end-to-end sovereign provision for the hardware, the software, the network access, the AI and Machine Learning capabilities, the security and the cyber resilience to achieve or to aspire to this destiny. It goes beyond addressing “just” data privacy concerns, and aims to stimulate significant national economic growth.

A recent report focused on the European digital sovereignty challenges estimates that as of 2020, 92% of the western world’s data is stored in the US, and that data footprint for the year was some 50 zettabytes—that’s 50 trillion gigabytes, and driven by such developments as significantly growing bandwidth capabilities for edge locations and IoT applications, this data footprint is growing 27% year-on-year. (As an aside, I had trouble envisaging 50 zettabytes of data ... so another way of thinking of it is if you wanted that much storage at home, then you’d want to connect up 5 billion X 10TB hard drives and have a very good home power supply.) That data footprint estimate would mean that by 2025, there would be 175 zettabytes of data being created!

Consideration of such a significant and growing data creation footprint and all the capabilities surrounding the data see the European Union (EU) looking to their increasing digital sovereignty capabilities to provide a GDP boost of at least 14% by 2030 across the EU—that’s representative of around 2 trillion Euros of GDP growth by 2030, which, as the report points out, is roughly the size of Italy’s current GDP to give some context on the significance of that growth resulting from digital sovereignty.

Legal and regulatory developments in this space are showing that there’s a growing global focus on sovereign digital capabilities, and whilst this is a much broader topic than data sovereignty and sovereign cloud, it serves to highlight the significance of the arena in which these topics play.
Global Data Privacy Arena

The global data privacy landscape continues to evolve, with some 145 out of 232 countries now having enacted data privacy laws of one description or another— that’s 62% of countries globally that have defined a local data privacy standard, up from 132 countries 12-months earlier in 2018-19. The COVID pandemic has far from paralysed developments in the area, with many countries continuing to uplift their existing data privacy laws, modernising and strengthening their requirements and largely holding the European Union’s General Data Protection Regulation (EU’s GDPR) up as the global gold standard when it comes to dealing with data privacy and controlling data and data flows for both public and private organisations and their respective citizens and customers.

Part of the drive to evolve the data privacy landscape goes towards the growing realisation of the immense social and economic value of sovereign data, whether collected and curated by public or private organisations. In particular, does a public body have the duty of care to proactively ensure that the economic research value of a sovereign data set, especially an aggregated data set, is fully realised? Not only in the direct economic sense, ensuring the availability of sovereign digital capabilities to process and extend the data set; but also with the indirect economic value that can be realised, enabling the full exploration and usability of the extrapolated data set, allowing for more meaningful decisions to be made from it and driving future social value. The potential reputational risk and opportunity cost for a public organisation not deriving sufficient economic benefit from national data, instead allowing policy to be determined by those with the loudest voice rather than based on best evidence, is an area for further discussion and only briefly touched on here. Who said the area of data was dull?!

The recent focus has not only been on the introduction and uplift of general data privacy laws, but also being mindful of economic and social value considerations, there is significant activity across the Globe in related areas, including on privacy aspects and dominance concerns of digital platforms, artificial intelligence (AI), infrastructure sovereignty and the proactivity of cyber protection mechanisms.

For example, just a few months ago in December 2021 in Australia, we saw the introduction of significant revisions to the Critical Infrastructure Act. These give the Australian Federal Government unprecedented powers to intervene in the security response of private organisations, and also see company directors being held personally accountable for a cyber breach, thus requiring more proactivity on their behalf than ever before when it comes to cyber security strategy and safeguarding the data assets of the organisation. The Critical Infrastructure Act (not to be abbreviated to CIA) has key implications across 11 core industry sectors: communications, data storage or processing, financial services and markets, water and sewerage, energy, healthcare and medical, higher education and research, food and grocery, transport, space technology, and the defence industry sector. The inclusion of these industries is a key callout by the Australian Government on the need for protecting sovereign capabilities across these critical areas, and ensuring a heightened focus on data protection and in turn data privacy.

Also in Australia there is current activity by the Government on reforming the Privacy Act, with formal submissions from interested parties closing recently. Following on from the earlier comments on viewing the GDPR as the gold standard on data privacy laws, the proposed reforms follow characteristics from both the GDPR as well as the California Consumer Privacy Act (CCPA). Recognising the key activities in the space, Australia also saw the introduction of the Certified Sovereign Data Centre Program with the release in January 2022 of the updated certification framework for the program. This framework emphasises the importance of hosting key government data in certified facilities “that meet enhanced privacy, sovereignty and security requirements”.

This Certified Sovereign Data Centre Program forms part of the Australian Government’s broader Hosting Certification Program and aims to ensure that certified facilities comply with data sovereignty, ownership certainty, supply chain risk management and transparent reporting and audit requirements through the application of a range of measures including technical risk management standards, security vetting of support personnel and proactive disclosure requirements. It applies to Data Centre Providers, Cloud Service Providers and Software-as-a-Service (SaaS) providers, and recognises that such Service Providers have a key responsibility in the stewardship of sensitive data with the application of more stringent controls beyond security and
focusing on the data guardianship and management of the data. It also captures recognition of the economic and social value of
the data, and of the ability of the Services Providers to further amplify the value of the data by ensuring efficient and
transparently defined services.

Activity across the Globe continues to have these common themes of sovereignty and privacy. The EU has just announced
agreement on the Digital Markets Act, with a focus on ensuring the “Big Tech companies” provide for fair competition to avoid
repercussions from their dominance of the market, as well as increased restrictions on how consumer data is utilised and
shared, noting that it largely leaves the data and citizen privacy aspects to the GDPR and rather addresses the economic and
liability consequences of exploiting personal and non-personal data. The Act now needs to be adopted by the European
Commission and Parliament. Similar laws are being considered in Australia, with the Australian Competition and Consumer
Commission (ACCC) in February 2022 releasing its Discussion Paper on “potential new rules for large digital platforms”. These
activities across Australia and the EU are also examples of nations and regions going beyond the recognition of the importance
of data sovereignty, and extending into the national or regional aspiration for digital sovereignty.

Both the EU Digital Markets Act and Australia’s proposed amendments by the ACCC have implications on the Artificial
Intelligence (AI) and Machine Learning (ML) capabilities that the major digital players utilise and how they gather, store and
process consumer data across their digital platforms. Regulations surrounding AI are still in their infancy, however if the
requirements outlined in these EU and Australian discussions are anything to go by, it is a signal that the privacy concerns of
such AI capabilities are certainly under careful consideration by these regulatory bodies. By extension, we can expect the
importance of country-specific or region-specific AI sovereign capabilities to be recognised as AI regulatory discussions
progress. Again, these considerations go beyond those of data sovereignty and highlight the importance of increasing the
overarching digital sovereignty capability. They recognise the increased social and economic value of having the sovereign
capabilities to aggregate, process, interpret and extrapolate base data sets through AI and ML and of the significantly magnified
value of such extended data sets. Further, they recognise the need to ensure this value is realised locally within the sovereign
jurisdiction of the nation or region, rather than having the value exploited by a foreign entity.

All in all, it has been and will continue to be a very busy privacy space, with many key and fundamental considerations
underway. However, despite the widespread implementation and strengthening of data privacy laws, much ambiguity and
uncertainty remains across the core data privacy landscape. Examples of the significant amount of interpretation and clarity yet
to be achieved include the slowness at which judgements as to the adequacy of a foreign nation’s data privacy laws by the
European Commission (EC) against the EU’s GDPR are being arrived at. As a result of this slowness, there is still considerable
uncertainty as to what the exact definition of “an adequate level of protection” is required to be when considering locating EU
data and the location for processing that data in foreign locations. With GDPR being viewed as the gold standard of data
privacy laws, these decisions have global implications.

We can see another example of the significant work still to be done in clearly defining data privacy when considering the
alternative means under the EU’s GDPR when it comes to judging the sufficiency of foreign data privacy regimes. This is
provided within Article 46 of the GDPR, where in the absence of an adequacy decision under Article 45, the transfer of
personal data may be undertaken if “appropriate safeguards” exist, however these are proving cumbersome and complex to
navigate. One only has to look to the Schrems I and Schrems II decisions and the judicial and expert dialogue around these
cases to just begin to appreciate this complexity, and what the implications for data relations between the EU and US and
ultimately globally as a result of this unfolding landscape.

Without surprise, there’s hugely divergent views on the delivery of the Schrems II decision. For example, Peter Swire, a US
national security and data privacy expert, commentator and official has said: “For national security experts, it is puzzling in the
extreme to think that citizens of one country have a right to review their intelligence files from other countries.” When we look
to what led Max Schrems to lodge his privacy complaint and then revise his complaint in the first place (we briefly touch on the
potential for indiscriminate data gathering by US Government programs both within the US jurisdiction as well as outside of US
Data Sovereignty and the VMware Sovereign Cloud Initiative

jurisdiction below), then personally I find it puzzling in the extreme that a citizen would not question the right of a foreign country that has no jurisdiction over them to be potentially gathering, collating and processing such personal data in the first place, especially when it is such all-encompassing vacuuming of data!

Perhaps what is just as important as the existence of the data privacy laws is also the question of the willingness and proactive enforcement of these laws—laws in and of themselves without the teeth to apply them will not achieve the desired impact. This is a highly relevant point when it comes to assessing the risk of where to place an organisation’s more sensitive data—the “crown jewels”, so to speak. If the data, or even elements of the data (for example, metadata or design artefacts), are to reside in a foreign jurisdiction that has some level of data privacy stipulation but has not shown the teeth or gumption to ensure that a citizen or consumer’s data privacy rights are enforced, then how would this be reflected in that organisation’s local reputation in the market or in the electorate, and what are the potential legal ramifications when it comes to ensuring the local data privacy requirements are met?

Specifically the ramifications for sharing security data across the Five Eyes Alliance is an interesting one to consider. For example, bilateral sharing under the allowances of the US CLOUD Act between the US and AU is far from finalised, but there are already significant security data sharing allowances made between the US, UK, Canada, AU and NZ under the Five Eyes arrangements. What does this mean when it comes to data privacy and the protection of a citizen’s or consumer’s data rights? This question goes beyond data sharing and data privacy between the US and AU, and certainly extends into post-Brexit UK data privacy considerations as to the ramifications there as well, and the interaction with the EU and globally.

Not All Data Is Created Equal

Not all data is created equal is a phrase with many connotations, however let’s briefly focus on the area of classifying data according to its sensitivity. Data can be sensitive for a variety of different reasons, for example it may contain identifying information of an individual, it may be a large accumulation of otherwise innocuous data that when considered in aggregate has significant worth or value, or it may have national security, national economic or commercial implications in the event the data was leaked.

A measure as to the sensitivity of the data can be the level of adverse impact on an individual or a public or private organisation in the event the data was to be maliciously or accidentally leaked, whether that be a commercial, national security, reputational or personal privacy impact. An example of a well-defined data security classification is that of the AU Government with the Protective Security Policy Framework (PSPF) defining sensitive and classified data, and indeed an organisation dealing with data of any potentially increased sensitivity will have their own similar definition.

It is data at these heightened levels of sensitivity, the “crown jewels” as we referenced earlier, whether that’s within a public or private organisation, where the key consideration of the sovereignty of this data comes into laser focus. This is the data that would be considered most critical and most sensitive, and in turn it is this data that must be best protected and controlled; the access to this data, the transit paths of this data, the processing of this data and AI/ML-based extrapolation of the data into extended data sets that must be well and carefully considered.

In looking at individual concerns when it comes to data classification versus those of national security, it is these two ethos that can be seen to have framed the different approaches globally to privacy laws and much of the resulting uncertainty and ambiguity that exists in the legal landscape today. We see the EU following a human rights and personal focus, with onus on the individual and the right of that individual to effectively “own” their data. Meanwhile, the US has followed a data privacy path with increased focus on national security and commercial concerns. It must be said that whilst these different approaches are not opposing and are indeed complementary, they are causing significant friction in the global arena. One has only to look to the ongoing dialogue between the EU and US due to GDPR ramifications to see the sparks from this friction, as well as ongoing legal tussles in the EU against the US tech-behemoths as they continue to fall foul of EU expectations, with the most recent example being the Digital Markets Act. However, when it comes to the simple consideration to actively protect the data in
question, no matter whether it's the “personal” or “individual” focus or that of national security and commercial interests, they are in agreement.

This difference in approaches to data privacy is evidenced with particular reference to the US Foreign Intelligence Surveillance Act (FISA) which enabled programs such as PRISM and UPSTREAM. These mass surveillance programs by the US NSA, revealed publicly by Edward Snowden in 2013, allowed unfettered collection of data across technology providers (Facebook, Google, Apple, …) and network cable systems. Whilst there were some legal exclusions, such as not collecting data from health care or pharmaceutical entities, amongst others, the technical limitations as to how the programs operated mean it is unlikely that data from these industries could avoid being collected and processed under such US surveillance regimes.

Once consideration is then given to the more recent US CLOUD Act (Clarifying Lawful Overseas Use of Data Act), which was adopted in 2019, in the context of other activities under FISA’s remit such as PRISM and UPSTREAM, then the possibility or risk of friction and conflict between competing data privacy laws is quite apparent. The CLOUD Act permits the US government to compel US-technology companies (both US and non-US companies that operate in the US and are therefore subject to US jurisdiction) to provide access to their data, no matter the location of that data, and even if that data is geographically located in a cloud environment that is under the legal jurisdiction of another country.

Sovereign Cloud

Sovereign cloud is not a new concept, however given the context of continued legal and compliance ambiguity and uncertainty in the global data privacy landscape, the relevancy of sovereign cloud is more significant now than ever.

The key aim of sovereign cloud is simple—namely, to better protect and to better control sensitive and critical data from both public and private organisations—ensuring data privacy, security and compliance for sensitive and regulated data and application workloads, whilst also controlling data transit flows and access to the data. The desired outcome is to better protect and secure citizen, customer and consumer data that may cause personal or commercial harm, reputational risk or general privacy concerns should it fall into malicious or unfriendly hands.

VMware Sovereign Cloud Initiative

Protecting and unlocking the value of critical national, corporate and personal data

Data Sovereignty and Jurisdictional Control

- **Autonomous legal operating entity** in the jurisdiction (i.e., no affiliate relationship)
- **Data is subject to the exclusive jurisdiction control** and authority of that jurisdiction
- **ALL data is resident** and controlled in that jurisdiction (e.g., customer data, meta data)
- **Foreign authorities** or legal entities are unable to assert authority over the data

*Figure 2: Sovereign Clouds—beyond data sovereignty and data privacy*
Data Sovereignty and the VMware Sovereign Cloud Initiative

It is a cloud environment that lends itself well not only to public organisations who collect and process data that have privacy and compliance considerations, but also very much to private organisations. In particular, those private organisations that operate in regulated sectors such as Health, Finance, Defence, Critical Infrastructure providers and similar areas have a high degree of relevancy for sovereign cloud capabilities.

To properly understand how sovereign cloud can be of significant benefit in protecting sensitive data, we should clearly define the different between sovereignty and residency when it comes to data and the hosting of that data in a cloud environment.

Data sovereignty and data residency are often combined and intermingled into a single confused statement. Ensuring data and the processing of that data sits within a defined geographical location for whatever reason (for example, policy, regulatory or performance reasons) is purely a matter of data residency; the idea that data is subject to the exclusive legal protections and jurisdiction of a nation is a matter of data sovereignty. Data sovereignty means maintaining authority and control of data within the nation’s jurisdictional boundaries.

To be clear, when considering the legal control of data, a jurisdictional boundary can be extended to encompass the limit of a legal entity, for example, referring to a grouping of political territories such as the European Union (EU) and legal bodies within the EU such as the European Commission (EC) and Court of Justice of the EU (CJEU) rather than just a member state such as Germany or France and their respective governments and judiciaries.

If an organisation has mission critical, private or regulated data in the cloud, what’s the potential risk and exposure of having it subject to a foreign nation’s laws? Keeping in mind that this question needs to be asked not only in relation to the core data set or workload under consideration, but also any artefacts associated with it. The question of artefacts is especially important when considering a public hyperscale cloud provider—what metadata is associated with the principal workload and where is this metadata located; where is the cloud support team located and what local security scrutiny and certifications are they subject to; what account information is maintained and what system is it housed in; where is identity and access management maintained; what design information is held and where is it located; where are backups or BCP copies held and are they purged as and when required?

Sovereign cloud allows an organisation to ensure they can confidently answer these questions without sacrificing any of the commercial benefits of cloud-at-scale and whilst maintaining all the flexibility, agility and visibility that we expect from a modern cloud environment.

As we touched on earlier, not all data is created equal, and it is important to note that a sovereign cloud readily co-exists with other cloud environments. Those other clouds may be on-premise, in a third-party data centre, at edge locations or with a public hyperscale provider with a mix of traditional or native cloud applications and services. With such co-existence, the “crown jewels” of the data workloads may be most suited to the sovereign cloud environment, whilst other cloud environments are leveraged for other data workloads, identifying where the Cloud Smart location is for each particular data workload, ideally subject to a common operating model to achieve the most efficiency in a true multi-cloud scenario.
Sovereign Clouds Augment Hyperscale Public Clouds
Leverage multi-cloud to share and monetise data while maintaining data sovereignty

Furthermore, the provider of a sovereign cloud will have a strong understanding of the local data privacy and protection landscape for specific data classifications and sensitivity, and therefore be well-placed to work with the organisation who owns the data on the Cloud Smart location for the data itself and for the applications that touch and extend the data.

The definition that an Australian sovereign cloud provider offers for their sovereign platform resonates very clearly in this space: “As a sovereign provider, AUCloud is owned, managed and operated in Australia. All services and data managed by AUCloud remain in Australia ALWAYS (including metadata, monitoring data and derived analytics data). All AUCloud services are monitored and operated in Australia by Australian citizens who have been security cleared to Australian Government standards.” Generalise this definition to be within any specific country, within any specific jurisdictional boundary, and we have a very powerful capability.

As we raised earlier, not all data needs to be treated equally—some data is certainly more sensitive and critical than other data, but when considering data that does have increased sensitivity (the crown jewels, as we referenced them earlier), what is the best way to manage the risk of malicious or unnecessary access to that data? Why should an organisation risk foreign jurisdiction over their data when they can efficiently and effectively have it protected locally in a sovereign cloud environment? What risk may an organisation be exposed to due to non-compliance with the ever changing local or regional data privacy laws? What are the expectations of an organisation’s consumers and citizens when it comes to the protection of their information?

The question we ask about where best such sensitive and critical data should be hosted and protected is a simple one: Where is the Cloud Smart location for this data? In this still evolving, ambiguous and uncertain global data privacy arena, the answer we look to is sovereign cloud.
Data Sovereignty and the VMware Sovereign Cloud Initiative

Where to Next?

Delving into data sovereignty and the resulting data privacy aspects is an interesting and crucial discussion when considering how best to protect and control the sensitive and critical data of a public or private organisation. Another aspect is the massive economic benefits of data, the exploding global data economy. When we look to this area in relation to data sovereignty, then the question becomes “Why should the value of a nation’s data be allowed to be exploited overseas?” Rather, the outcome should be with the data investment being enjoyed locally to drive and stimulate the local economy. “Data is the new oil,” says Clive Humby, British mathematician and data scientist—and he most assuredly is correct, and I look forward to exploring this aspect a little further shortly!

References

2. GDPR, Article 45: https://gdpr.eu/article-45-adequacy-decision-personal-data-transfer/
9. Max Schrems, Schrems II, Background: https://en.wikipedia.org/wiki/Max_Schrems#Schrems_II
12. Australian Proposed Legal Reforms, Privacy Act: https://parlinfo.aph.gov.au/parlInfo/search/display/display.w3p;query=Id%3A%22media%2Fpressrel%2F6577790%22;src1=sm1