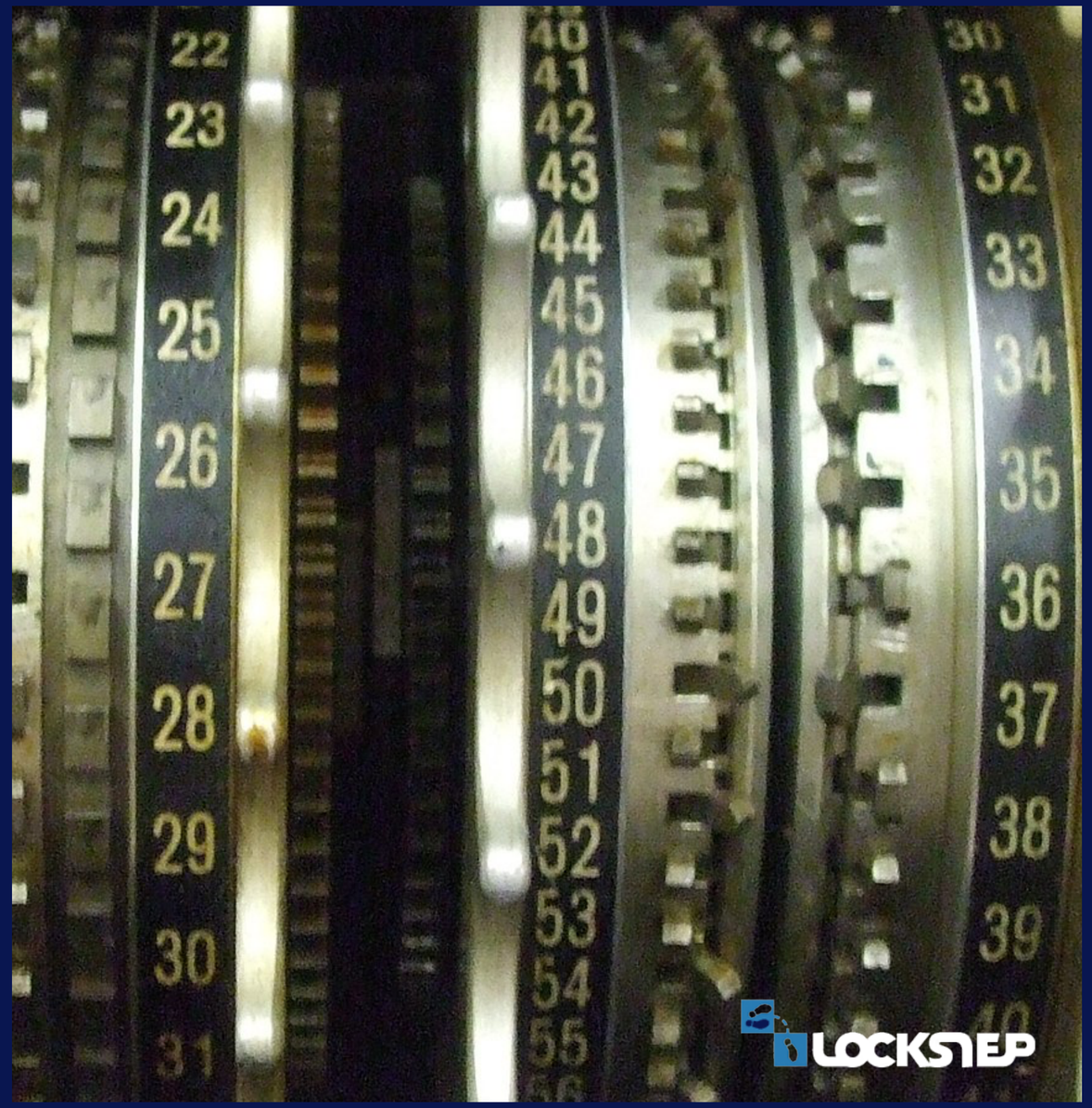
Stephen Wilson, Lockstep Consulting

How a data protection infostructure can safeguard value chains

April 2023

Adapted from a presentation given at CyberCon Canberra in March 2023. Photo: OS/WikiMedia Commons (CC BY-SA 4.0).



Data breaches continue to plague us

Data-sharing is now a social and economic imperative, yet the risks have become obvious.

A global data-sharing marketplace is within reach, where consumers and businesses will access quality-assured information from any digital application. Data can be made as safe and reliable as water, gasoline, and electricity.

Part of the solution is to extend verifiable credentials such as mobile driver licences from certifying facts about people to *certifying any data about anything*.

Digital wallets can provide the same user experience we enjoy with payment cards. Just as we “Click to Pay”, we should be able to “Click to Prove” any verified data.

But data wallets need networks to bring users and providers of data together in an equitable two-sided market. Wallets plus business networks equals data protection *infostructure*.

Optus showed us the risk

In late 2022 the Optus data breach exposed to criminals the identifying data of half the Australian adult population.

Within days, some state governments fast-tracked the means to replace digital driver licences where applicable. That was an inspired, digitally agile response. But all the same, for years to come, valuable data will be on the black market.

What is the best long-term response to the Optus incident? These breaches are multidimensional events, with many pertinent lessons. Yet one strategic factor is not being talked about.

How do we make stolen data useless?

Paradigm shift

Paradigm: “Universally recognised scientific achievements that, for a time, provide model problems and solutions for a community of practitioners.”

– Thomas Kuhn, *The Structure of Scientific Revolutions*

Data breaches are not an identity problem but a data problem

The way we respond to problems is determined by the way we frame them.

Technologists often use “paradigm” to mean model solutions, but paradigms are really more about model problems. The way we look at the world and the way we frame our problems are frozen for long stretches of time.

Data breaches do not point to an identity problem but rather a data problem. There are pinch points throughout e-business where important data is vulnerable to substitution. All fraud boils down to unreliable data.

Digital identity practice is making a genuine paradigm shift from *who* to *what*.

The most important developments — such as the FIDO Alliance and verifiable credentials — are not about “identity” and don’t even use the word.

We rarely need to know who a counterparty really is.

Instead, we need to know certain specifics, depending on the context.

We only got into the bad habit of hoarding personal information because we lost confidence in the core facts that are central to transactions, such as credit card numbers, or age, or eligibility for Medicare rebates.

We should think differently

Instead of identity, we should ask:

- What do we really need to know about the counterparty?
- Where will we get the data?
- How will we know that it’s true, or fit for our purpose?

C2PA

GO FAIR

Consumer Data Right

ONDC

Recognise the importance of data

The World Bank in 2021 called for "a new social contract for data to enable the use and reuse of data to create economic and social value promote equitable opportunities to benefit from data, and fosters citizens' trust".

There are many examples of a fresh focus on data and data quality:

- Coalition for Content Provenance and Authenticity (C2PA) of Microsoft Research, the BBC, and the *New York* Times.
- Governments mandating access to publicly funded research; FAIR Principles (Findability, Accessibility, Interoperability, and Reuse).
- Australia's Consumer Data Right (CDR).
- And more recently, Australia's *Data Availability and Transparency Act (2022)* with its DATA scheme.

"Forging a new social contract for data is a pressing domestic policy priority that will require strengthening national data systems and engaging all stakeholders at the national level." — The World Bank

Contain and protect data

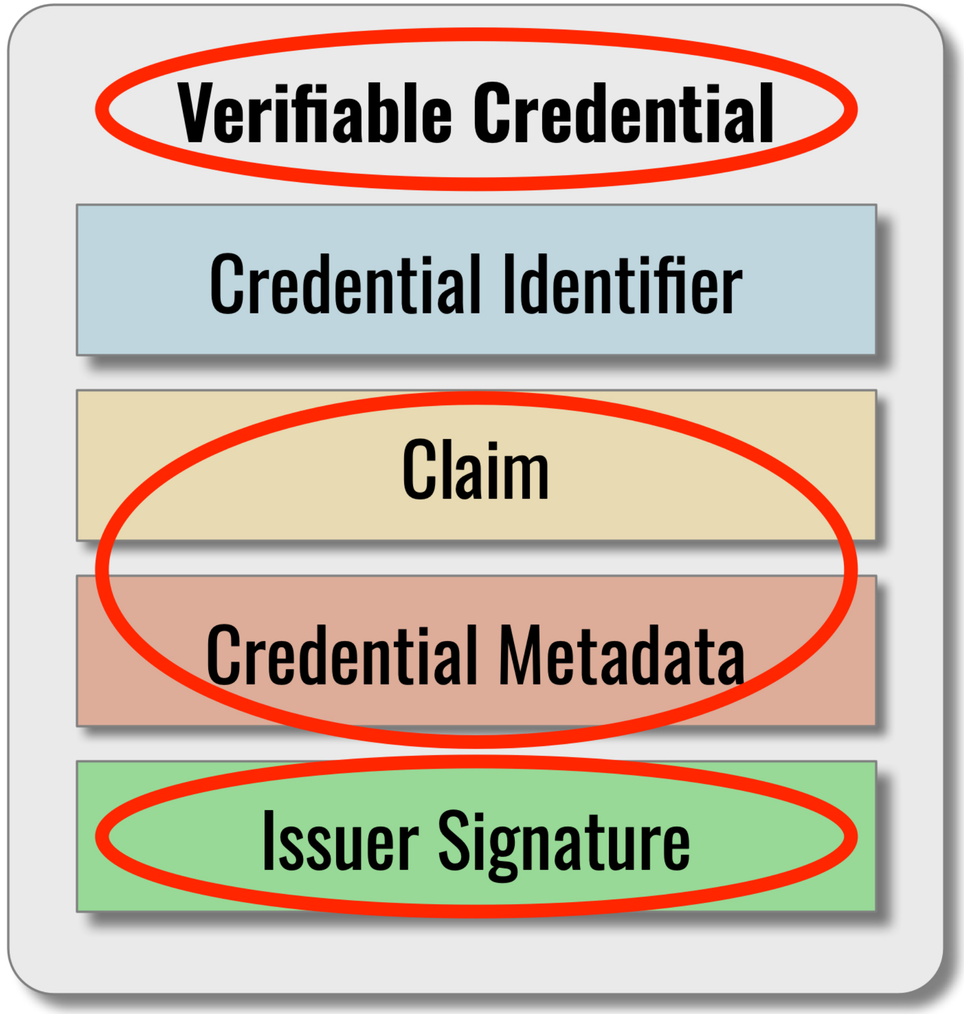
Verifiable credentials are becoming standard containers for claims vouched for by recognised authorities.

Essentially a verified credential holds verified facts about a credential holder and binds them together in a data structure cryptographically signed by the issuer. Usually a key pair of the credential holder is included. As a result, when the holder signs a transaction using that key pair, the claims and the authority of the issuer are bound to the transaction too, and go wherever the signed data goes.

Verifiable credentials can hold additional metadata about holders and their claims, and protect the value of data more broadly than focusing on breaches.

The currently best-supported verifiable credential data model comes from the World Wide Web Consortium (W3C). Its data model captures the design thinking we've described here for authenticating any important data:

- What do you need to know? *Claims and metadata*.
- Where will you get the data? *Verifiable credentials*.
- How will you know that it's true? *The issuer's signature*.



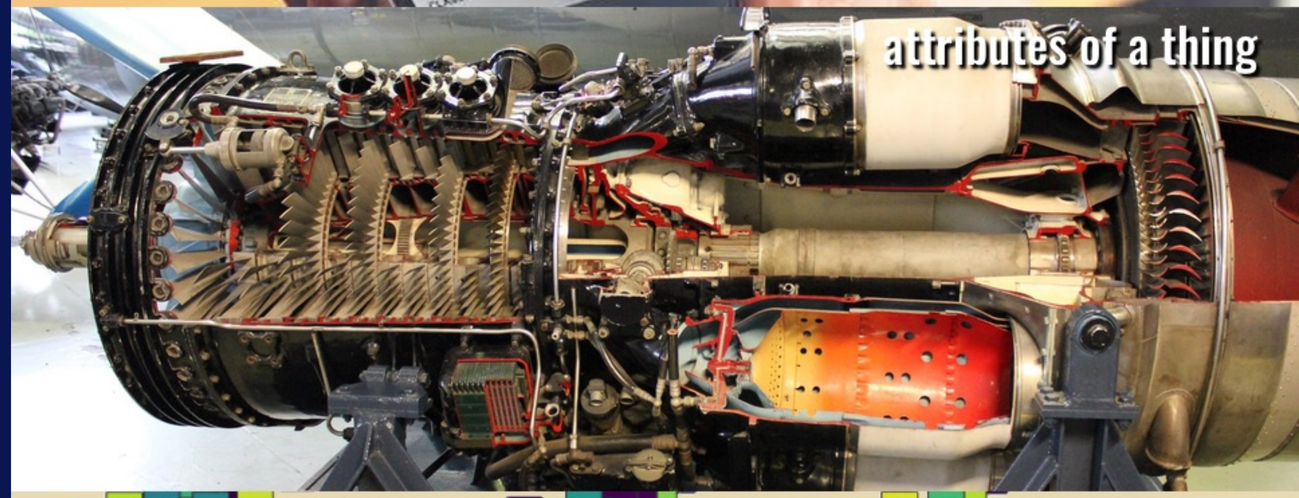
W3C Verifiable Credential Data Model 1.0

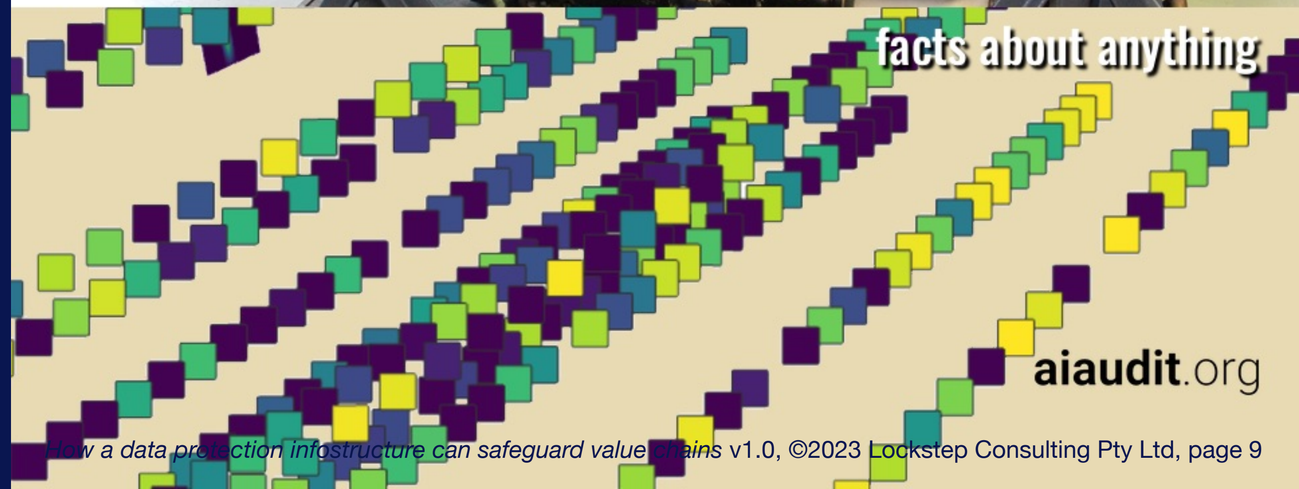
Extend verifiable credentials to all types of data

The digital identity industry has given us ideal tools for data protection writ large.

We can fix the originality, fidelity, and authenticity of all data. We can make stolen data useless to criminals. We can tell fakes from the real thing. And we can ensure the multifaceted quality of data as its makes its way through information value chains.

To build data protection into the digital economy, the first step is to extend verifiable credentials tools.



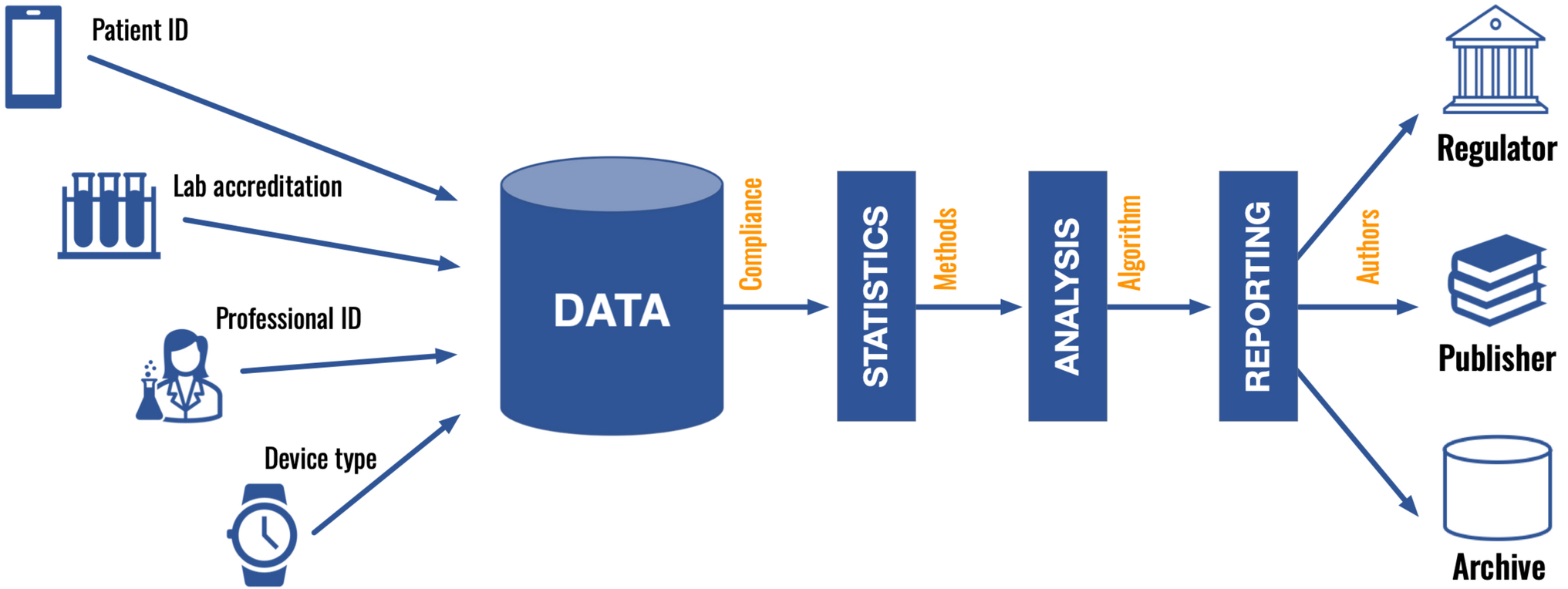


Let’s organise the data supply chains

Data supply chains, also known as information value chains, have been forming for years in the digital economy, on an ad hoc basis.

The time has come to make these supply chains more orderly, secure and transparent, by systematising the qualities and metadata that makes data valuable.

Example: Clinical trial data flow



Consider how data's value grows as it moves through the course of a clinical trial, and the qualities that matter to the respective stakeholders.

Raw data is collected from a variety of sources, including devices, patients, investigators, and laboratories, and passes through successive processing steps. Different parties need assurance of device characteristics, correct patient enrolment, the professional standing of investigators, laboratory accreditation, security compliance, statistical methods and algorithms, and authorship.

Couple supply and demand

We've looked at the supply side for the data economy. Now we need to bring in the demand side to create what economists call a two-sided market.

The paragon of two-sided markets is the global credit card system. Merchants happily accept credit cards from complete strangers. Merchants don't even recognise the issuing bank behind the shopper's card. How is it that a merchant can rely on credentials from far-flung issuers without any bilateral arrangements and without any legislation?

The key is the merchant's bank, known in the card industry as the "acquiring bank". The acquirer onboards merchants into the scheme, provides terminal equipment or APIs and gateway servers for accepting cards, and above all, signs the merchant up to a services contract, the form of which is highly uniform worldwide.

Infostructure

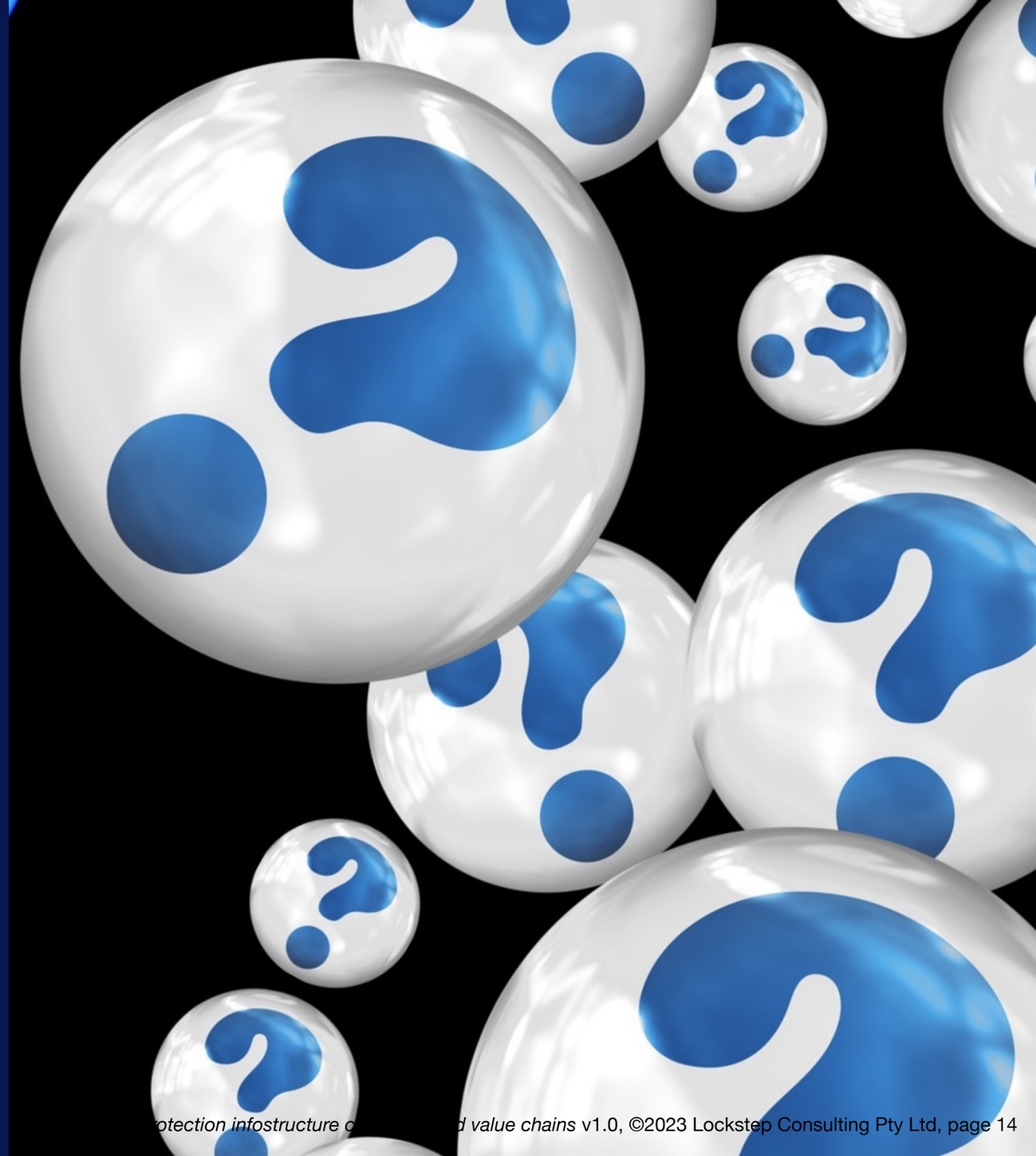
“An organisational structure for the collection and distribution of information… hardware, networks, applications, etc. used by a society, business, or other group.”

— *Oxford English Dictionary*

Verify all the data

The value of a piece of data can originate in many ways, and can be conveyed by metadata that verifies its source, history, compliance, authorship, and so on.

All important data has a story, and we can tell that story through metadata bound to the data.



Verifying all the data requires an intermediating platform

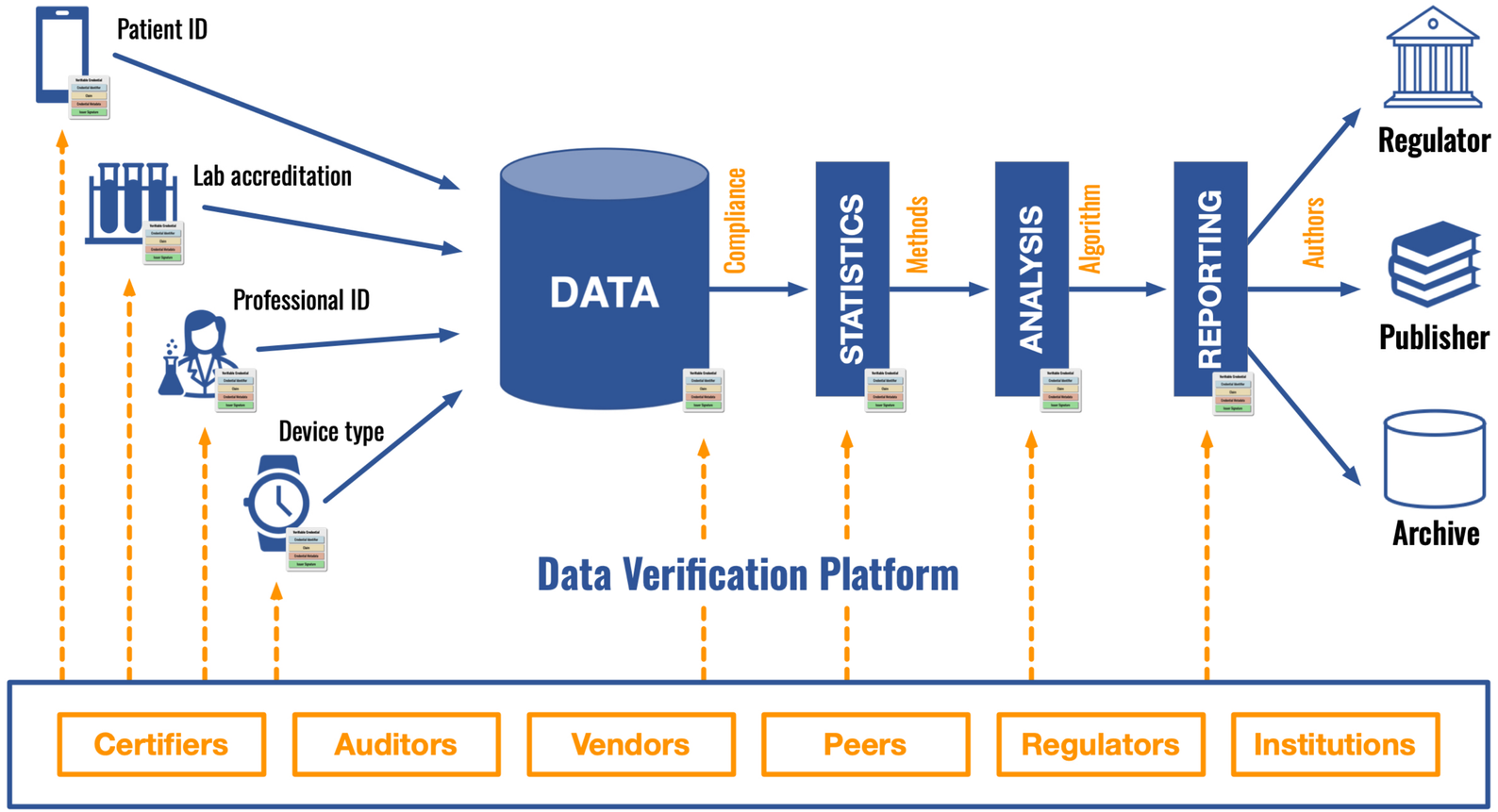
Verifiable credentials can be issued to non-human subjects, to prove any data about anything.

To be able to accept verifiable credentials and make sense of them, risk owners need to know what verifiable credentials to expect, which issuers have been approved, what each of the credential fields mean, and so on.

They also need copies of the root keys with which to verify signatures.

All this information must be available without needing bilateral arrangements between every issuer and every risk owner.

For the acceptance of credentials to scale, as with card payments, an intermediating platform is needed.



Everyday data verification is becoming bread and butter

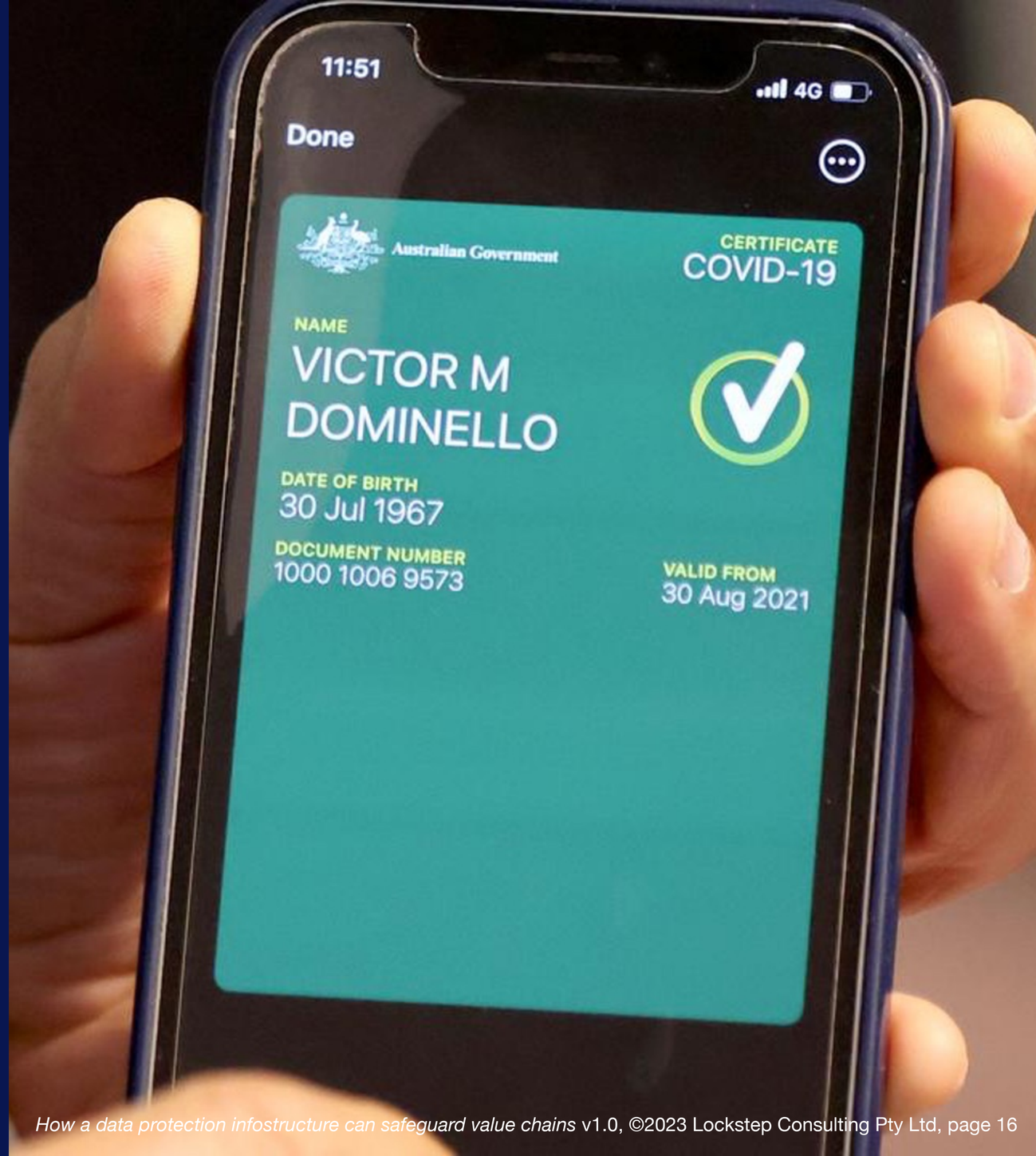
The clinical trial example highlights the importance and complexity of data-sharing in highly technical data supply chains, but everyday data verification is in fact even more urgent.

It is increasingly clear that the efficiency dividends of digitisation depend on better verification processes online.

The root cause of all identity fraud is unreliable data.

That is, we need better data, everywhere.

- Account opening and KYC
- Proof of Age – cross border
- Proof of Age – for children
- Medicare provider fraud
- Push payments payee verification.



Security, privacy, and ease

A globally scalable data verification platform looks like a heavy lift, but this sort of infostructure has been built before — and recently!

In the last few years we have seen mobile phones and wearable devices extended to hold secure boarding passes, payments credentials, even hotel room keys. Soon, COVID vaccination certificates, mobile driver licences, and the European Union electronic ID will also be available too in the standard mobile wallets.

We have the tools — data networks and cryptographic endpoints — for proving the origin and integrity of special packets of data. And new data-sharing rules are being set by governments worldwide.

People will soon be able to prove any facts about themselves, with the same security, privacy and ease of use that they enjoy with “Click to Pay”.

Let’s extend the infostructure so everyone can Click to Prove.

Stephen Wilson, Lockstep Consulting
lockstep.com.au
@Steve_Lockstep