The Decentralised Identity Opportunity

Web3 Labs

Contents:

Introduction

Decentralised identity is a new and emerging technology that solves many of the challenges with online identity today. Decentralised identity, a core tenet of Web3, revolutionises the way we can interact with systems and transact with both businesses and end-users.

Decentralised identity allows for better and easier interoperability across otherwise disconnected functions. It does this while keeping the user in control, with better regulatory compliance and stronger guarantees around privacy. With an expected CAGR of over 85% and estimated to be a multi-trillion dollar opportunity by 2030, now is a good time to begin exploring the opportunities available.

Through this technology, we can solve password fatigue and provide a better user experience by enabling the user to interact with less concern while maintaining and improving on best practices for account security. However, what decentralised identity enables goes far beyond just making it easier to sign up and sign in to applications. With this technology, we can open up new data-sharing ecosystems and unlock value that was previously locked into data silos.

Enabling trust online isn't limited to just humans interacting with online services, it can also be applied between organisations and between connected devices. It is an enabler for trust between any type of peer-to-peer connection, something which helps unlock new business opportunities and improves existing business relationships. This brings new revenue streams to otherwise underutilised assets.

Key to this ability is the use of blockchain technology. Fundamentally, decentralised identity utilises blockchain technology in such a way that no private data is at risk of ending up on a public blockchain. This is essential if we are to remain compliant with regulations like GDPR and CCPA, because as we know, once something is on the blockchain, it's there forever. This is also key in a B2B setup where relationships and related processes must be guarded.

The following chapters dive into challenges and issues with existing identity solutions and explain how decentralised identity can help fix, improve and expand on these. The content is not meant to be overly technical, but rather to introduce key concepts and areas within the scope of decentralised identity. We hope this helps you to identify potential opportunities for you and your business, and acts as a catalyst for further exploration.

[1] https://www.marketsandmarkets.com/Market-Reports/decentralized-identity-market-59374755.html
https://www.polarismarketresearch.com/industry-analysis/decentralized-identity-market

The problem with online identity

Have you ever signed up for a bank account online? Or maybe a brokerage account? Maybe with a challenger bank, one with no physical branches?

If you have, you've likely been through an onboarding process that involved you scanning your passport or driver's licence, perhaps even taking a selfie. Maybe you had to attach the scanned JPEG image to an email and send it. Similarly, the gold standard of proving you reside at your place of residence is providing a PDF version of a utility bill.

And that's trying to prove something from the point of view of an individual consumer. What about the receiver of the above information? What awkward manual data entry routines and checks do they need to make use of this information? And how reliable is it? Is the quality of this data even quantifiable?

The internet, and hence our online presence, started in earnest in the 1990s. Yet in the first half of the 2020s, 30 years later, the adage "on the Internet, nobody knows you're a dog"[2] still remains true due to the ease with which anyone can provide data on behalf of someone else.

It's not that we haven't tried. Tools that allow people to sign documents electronically have existed for over 30 years. Whilst we can argue about the ease of use on these, it's not really a smooth user experience that's been the missing part.

At the heart of all of these problems has been a lack of trust in who is actually on each side of a transaction.

While the internet enables peer-to-peer communications, without the need for any particular intermediary, the services we've built on top very much represent intermediaries. These intermediaries, giants of the Web2 era, like Apple, Google, and Meta, to name a few, have, for the lack of a better alternative, been entrusted as online identity intermediaries. Or, as is the case in many countries, the government has established some online mechanism for its citizens to electronically sign in, with the government acting as the trusted third party.

While this is better than nothing, it puts tremendous responsibility and power in the hands of a few actors. These actors not only gain the power to attest to your sign-in request, but they also gain insights and data about how and where you utilise this.

[2] https://en.wikipedia.org/wiki/On_the_Internet,_nobody_knows_you%27re_a_dog

Maybe it isn't a big problem that Google knows when you sign in to some third-party service, or that Meta knows what bank you use and when you access it .. no, wait .. that is a big problem!

The first obvious problem is, why should they need to know this? Just because it's convenient for you to not create yet another profile with yet another password? It's a broken system when a federated solution, leaking behavioural information, is assumed acceptable.

Secondly, because federated systems concentrate power in single entities, you can only attest so much value to that before you run the risk of too much exposure. Take an example: You want to start an online payday lending business. If you let people sign in to the system to request a loan using Google's single sign-on (SSO), would you feel comfortable with Google claiming a particular sign-in request is from a particular user? The federated system can easily impersonate any one of its registered users, so it wouldn't be advisable to trust this request beyond a certain low level of exposure. And how much KYC does Google do anyway?

As a user, what happens if you break some T&C of Google, and they block your account? Not only do you lose access to all Google services, but you can't even sign in to other unrelated third-party services that you used SSO to sign in to. There are ways to re-enable access to these third-party services, but it's rarely straightforward.

Hence, end users are stuck with either creating many individual profiles at different services, without an easy and privacy-preserving way of linking these, or picking one of a few major federated SSOs, leaking information to these with the risk of being impersonated and censored at any point in the future.

It's important to realise that this isn't just a problem for the end user. Without trusted metrics derived from trusted data across many entities, organisations and devices, we have to add bureaucracy on otherwise productive relationships simply because we don't have a clear risk profile on someone or something. Without a better and more up-to-date view of how your upstream and downstream business partners are currently operating, the process introduces lag and uncertainty. This reduces transaction volumes and leaves money on the table. In short, we could have done more together if we only knew a bit more about each other.

Like internet-enabled peer-to-peer communications, we need a system that can enable peer-to-peer trust, so that we can establish credible connections without selling out to any intermediary. In this ebook, we will cover how such a decentralised system can be realised, building on top of open standards and technology, and how this can enable and unlock new and improved business models.

Data is a liability

Before we dive into how we can solve online identity, we need to talk about data. Data is the oil of the digital economy. If you have data, you have something of value. The more data, the more value you could extract. This brought us major players like Google and Meta, companies that would "give away for free" products and services. In reality, users were the product. Their behaviour, as represented by data, would be distilled and bottled, before being sold to the highest bidder.

This is of course, good for them and also helps explain why they are happy to act as SSO solutions, as it gives them even more data. But, as we've touched on, not always so good for end users.

Over time, the regulatory landscape caught up and put in place rules and legislation that turns data into a liability. Through acts like GDPR, CCPA, and others, companies like Google and Meta can't just harvest unlimited amounts of data and do whatever they want with this. They have to give you a choice and guarantee that data about users isn't used beyond a certain scope.

They could stop providing free services and products. However, as harvesting their users' data is their core revenue stream, it's better if they can find a way to still monetise this data within the regulatory framework. This isn't easy for many of these companies, with Meta putting its hand up saying they have no idea where all this data flows.[3]

It's a big job to change the enterprise architecture in such a way that it works for everyone involved. Modern online identity solutions, known as decentralised identity, provide exactly this type of architecture. This allows us to ensure that data is shared in a compliant manner, by putting the user, or more correctly, the holder of data, in control.

Not only does this allow existing data-driven businesses, like Google and Meta, to find new and better ways of selling data, it also enables companies that currently have valuable unused data to begin releasing revenue on this. This is because ***in a decentralised model, where peer-to-peer trust can be established, actors that previously were left behind in the Web2 world of mammoths, can now efficiently play a role in the open data economy.***

Making data a liability is therefore an opportunity and a driver to build a more trustworthy setup. One which will enable any type of transaction, be that large or small, frequent or infrequent, that allows both large and small businesses to participate on equal terms. It will put users in control, provide for more trust, and improve risk and compliance clarity.

[3] https://www.vice.com/en/article/akvmke/facebook-doesnt-know-what-it-does-with-your-data-or-where-it-goes

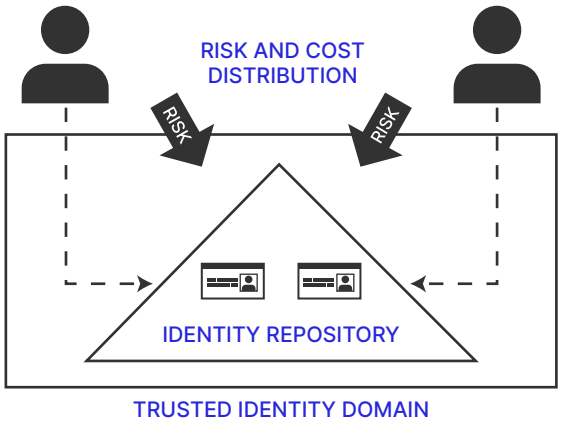
De-risking data with decentralised identity

Before we can successfully implement a decentralised identity solution, we need to know what types of risks we want to avoid, reduce or eliminate. We've already mentioned a few, but let's go through them more methodically.

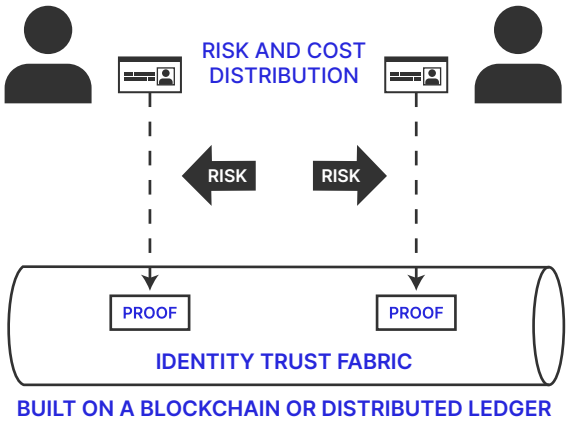
Federated identity solutions allow for profile data to be stored with one central provider. The convenience here is that a single identity can be reused across multiple systems. This saves the user from creating many profiles and often provides operational benefits as we get to reuse infrastructure, technology, and operational teams.

The problem here is that, as more and more systems start depending on this centralised identity solution, it becomes an ever-increasing risk factor. One of these risks being that, as a centralised solution, the implications of any downtime become noticeable across a vast array of dependencies. The other side of centralised systems is that as the identity solution grows in scope, it becomes an ever-increasing target for hackers. It's the very opposite of diversification, instead it puts all your eggs in one basket.

Centralised Identity



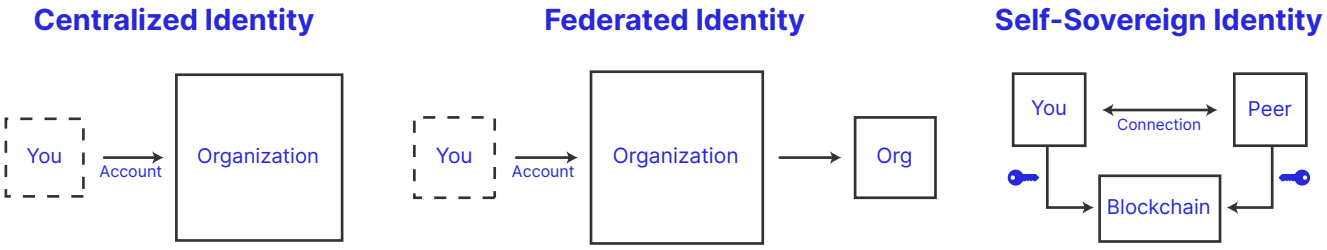
Decentralised Identity



While hackers attacking a centralised solution with an ever-growing vigour is troubling enough, we also risk insiders impersonating a user. When using a federated solution, you inherently trust what the platform claims about a session. If it claims that this is Alice, you'll have to accept that.

It's clear then that we could reduce these types of risks by avoiding centralisation. If the identity of users is spread out, not only does it become a less desirable target, but it also reduces the impact of failures. Decentralised identity takes this all the way, by letting the user, device, or entity create their own identity. Any such user of a decentralised identity solution would create their own public identity backed by a private encryption key. They keep their private key private.

The process of creating a public/private key pair is nothing new, but how we distribute the public key is new. While we previously had to delegate trust to a centralised or federated solution, hence introducing a third-party, when uploading our public key to them, with blockchain technology we no longer have that problem. We can now distribute public keys securely to the whole world, and everyone who receives this public key can be sure of its integrity as it is secured by the blockchain it resides on.



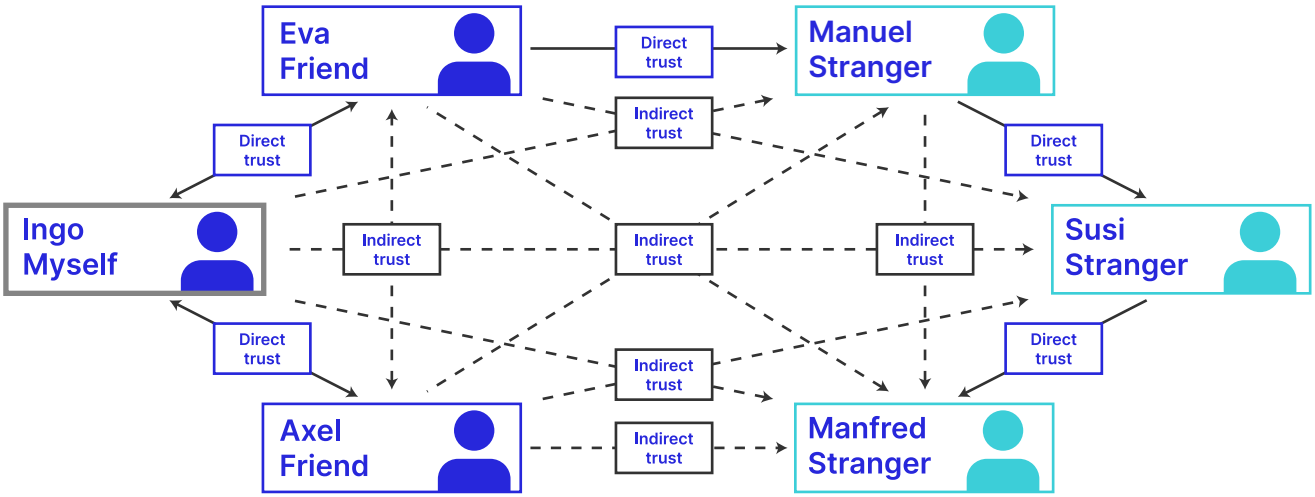
This way of having decentralised identities, through public/private key pairs, solves the centralisation problem. It also helps solve the inevitable integration between two or more federated solutions. Federated solutions have no good explanation for how to integrate two or more of them, other than suggesting expanding the union. All systems that want to support two federations would individually need to support both. Hence, federated solutions are making the risks even worse, by increasing the centralisation and blast radius even further. Or, it's not really functioning as a federated solution, but instead as two unrelated sources.

Federated identity isn't scalable, it just happens to be convenient up to a certain size.

Instead, in a decentralised setup, we let all users go ahead and create their own identity, or "bring your own identity" (BYOID), just like we let users bring their own devices when accessing your systems. How can we know that a particular user is who they claim to be? Haven't we just introduced another risk here?

To answer this, we need to start looking at the big picture. Firstly, we will need to let the user associate their new decentralised ID with certain claims so that we can start to trust that when user X claims this is their ID, we have something connecting that to something tangible. How can we do this?

We could allow Bob to attest that Alice, as identified by her self-generated ID, is in fact Alice and not Charlie. While this doesn't always hold much weight on its own, it is often the first seed towards building a "web of trust". What if a hundred people or companies attest that this is Alice? We can recursively iterate on this to determine the overall trustworthiness of this network and its claim.



As we start building a bigger trust web, we start having quantifiable metrics on how trustworthy a particular claim is. If a government agency, identified through a well-known public ID, attests to the name, date of birth, address, and other such information, we can likely attribute a high level of certainty to this claim. And if more than one well-known and trustworthy actor like that attests to the same, the quality of the claim improves even further.

We can see how this is building a system that becomes more robust as we keep on growing it. Each time we build a trusted relationship with a new actor, we can learn more about the trustworthiness of others, by analysing their relationships.

Finally, let's touch on the regulatory risks, with data now also being a liability as previously stated. It was mentioned that not knowing how and with whom we share personal information is a big problem for many organisations, with GDPR, CCPA, and related regulations requiring organisations to manage this. What if you let the subject, i.e. the person identified by some data, obtain ownership of this data and share it with others as they see fit? Clearly, at that point you're off the hook, since you're no longer directly involved.

Decentralised identity solutions allow for exactly this, where the user, also known as the subject or holder, is kept in the loop when information you hold about them is shared with others. They are in control, but as we'll describe in the following chapter, it does not mean business opportunities become irrelevant. Consequently, we have a responsible, compliant, and scalable way of sharing data between organisations, departments, or other such entities, with quantifiable metrics around the risk profile of those involved.

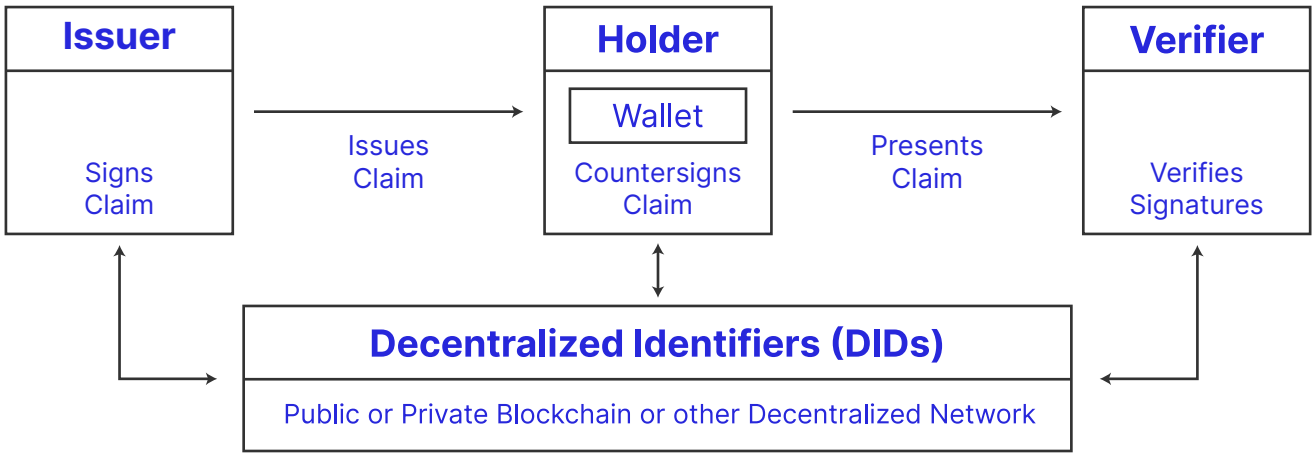
The new business of credential issuance

While at the start of this ebook you would have been forgiven for thinking it was mainly about usernames and passwords, and how we could improve on that situation with decentralised identity. But, it should now be becoming clear that this is about something bigger.

Once we've solved the centralisation and integration risks and challenges with existing fragmented identity solutions, something decentralised identity does by letting users "BYOID", we can quickly turn our focus towards how we can leverage these new abilities. Here, having others attest to claims about someone or something creates new business models. Some might call this reusable KYC, and others a decentralised "web of trust''. Both are valid descriptions, if not limited in their imagination.

While the user sits in the middle, in control of the flow of information about them, on each side we find an issuer and a verifier of this information. Both of them are themselves identified through their own public and private key pairs, and as such are just actors otherwise indistinguishable from anyone else. But they take on certain roles in this sequence of events.

DIDs enable digitally signed varifiable claims



The verifier is interested in obtaining some kind of information about the user and has been granted access to this information from the user. The verifier would then cryptographically verify (hence the name) the integrity of this information. This means checking that it comes from the expected issuer and that it involves the expected subjects. It is clear then that certain issuers have more favourable terms applied to them than others, as those entities with existing high levels of trust would have a head start when attesting to a claim.

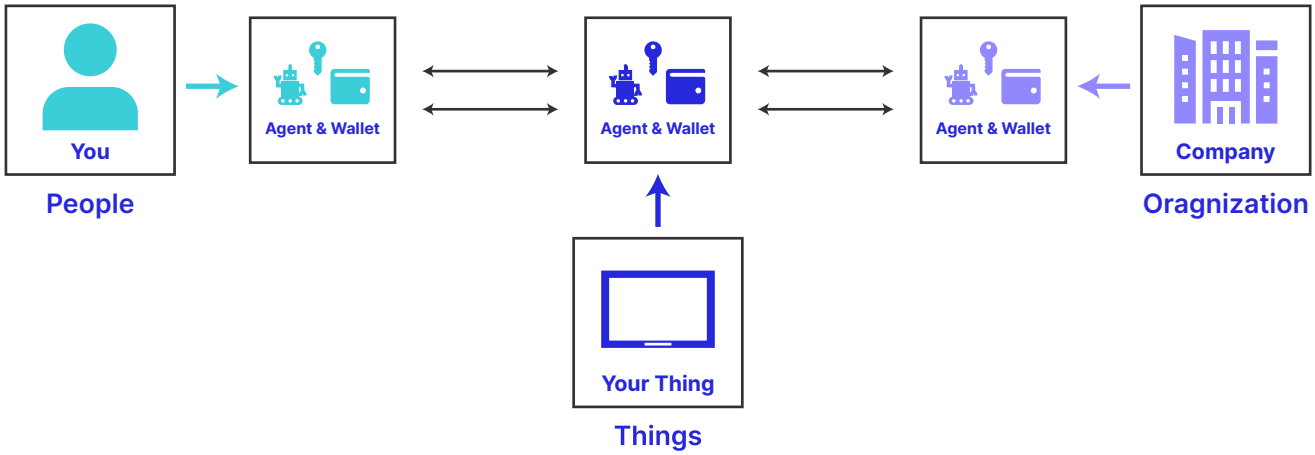
Among entities with a high level of existing trust, we often find official government agencies. This can be in the form of a driver's licence or passport, each holding a high level of authority in most societies. And while it would bootstrap the decentralised identity ecosystem to have these agencies issue digital certificates and credentials, it doesn't have to start or stop with them.

With a quantifiable risk profile on those involved, we can also start accepting claims by private enterprises, like telcos and banks for example, companies that have already invested heavily in KYC. And because the information flow ensures safe and correct sharing of private data, previously untapped but highly valuable data can be commercialised.

For example, let's go back to the initial description of how you maybe had to once attach a scanned PDF copy of your passport or utility bill to an email to complete a KYC process with a broker. With decentralised identity, you'd instead have your telco electronically issue you a proof of address, not as a PDF, but as what's known as a Verifiable Credential. This has been cryptographically signed by them and you and resides with you in a wallet application on your phone. You then simply open this app and security share this information with the broker, all online. This is easier for you and also more secure for everyone involved.

While we've said that users gain control of their data, like the example above, it does not necessarily mean we, as the issuer or author of said data, have given up commercial interests in it. As the originator of this data, we can participate in governance efforts that allow us to ensure commercial interests are maintained between issuer and verifier.

A licence can be attached to the data. The verifier knows who issued the data, as the data was cryptographically signed by the issuer. Also, through related mechanisms in decentralised identity solutions, a peer-to-peer connection can be established between the issuer and verifier.



This allows for a decentralised method of ensuring commercial interests are maintained, again without needing to introduce any trusted third parties. Because it's all machine-readable, quantifiable, and automated, these microtransactions become economical. Data that was costly and tedious to obtain and maintain can start paying dividends.

Utilising untapped data silos

The flow of information and data in decentralised identity solutions are clearly defined and often described within what's known as self-sovereign identity patterns. Here, as stated earlier, data control and sometimes ownership is in the hands of the user, also known as the holder.

It might be helpful to relate this model to what we have in the physical world. Take passports as an example. These are issued by a specific authority and given to a holder who is usually the subject, i.e. the person identified by the passport. The passport itself contains the information, and the holder is free to show the passport to anyone they choose. Usually, they would show it to someone when travelling and wanting to enter a foreign country, or returning home after a trip. But it can also be used to prove identity in other circumstances, like when obtaining a job, and so on.

When sharing the passport with someone, the receiver is free to consume the data on the passport without informing the issuer about this. But we can also imagine models where the consumer, or verifier if you want, reaches out to the issuer for additional information. This does inform the issuer that data has been shared with the verifier, but only after the fact that the holder initiated and confirmed this sharing request. This gives us certain guarantees around the sharing process and what steps were taken previously, allowing us to stay within regulatory boundaries.

An example of such additional information could be a credit score. Let's continue our earlier telco example, providing your customers with a credential that allows them to prove their name and address, based on your KYC data. As additional information, you could provide to the verifier your internal credit score of the customer, based on their payment history. That would be an add-on service, shared directly between the telco and the verifier, but only after the customer explicitly initiated the flow first.

With this in mind, we can formulate a clear description of the steps we need to take to make use of untapped data silos. An untapped data silo is any kind of database, holding data that can be of value to a verifier, and hence optionally release new revenue streams. We've already mentioned KYC data and payment history data, but we could also mention various types of preference data obtained from hotels, car rental companies, online services, etc. It could also be health-related data, proof of employment, proof of insurance, membership or subscription confirmations, your payslip, and so on.

You are not in this flow giving holders or verifiers direct access to your data silo. Instead, you are extracting predefined attributes from the silo, packaging this up into its own data structure, signed by you and others involved, before being given to the holder. It flows from you to the receiver through a one-way street, and this means we can tightly control access rights and data-sharing processes, reducing unauthorised access risks.

While some types of data can be useful in many places, it can also have a very narrow use case and scope, like perhaps an invoice or purchase order, something that isn't likely to be shared much beyond any particular supplier or customer relationship. In these cases, the ecosystem value might be in helping to define relevant schemas, adding structure to the blob. That is, team up with others and push on with a standard for what data fields should be included.

That can be a daunting task, but being able to agree on a data format can unlock value previously stuck behind manual processing and uncertainty. It can help optimise a business process, automating it for the benefit of everyone involved, and it can start at a small scale for quick turnaround if need be.

Identify the right opportunity

In our description of decentralised identity and the flow of data it enables, we've usually included humans as users, sitting between issuers and verifiers. But the system is open-ended, and we can easily consider devices such as phones, sim cards, bank cards, cars, and so on, as subjects, holders, and users, or humans as issuers and verifiers, in this flow. Any combination is possible here, and therefore, a broad potential of opportunities exists. It can also be that on each side of a connection we find agents representing a business unit, useful when suppliers and clients connect with your business as part of everyday business transactions.

While we expect a network effect to take hold and quickly accelerate the decentralised identity opportunities, it is often helpful to start with concrete application-specific use cases. An application-specific use case often reduces the number of moving parts, allowing us to lock in the various actors in the flow. This makes it easier to manage the project and serves as a good starting point.

When identifying the right opportunity in your business, you might even keep it all internal, where systems under your control start using decentralised identity technology as a way to provide better risk management around data flow and provenance. Because decentralised identity solutions are at the core decentralised, as you get your internal system running, it might serve as a good stepping stone towards including external partners and customers when the opportunity arises.

It's important to avoid approaching this with a "Web2 mindset". The aim here is not to build another centralised platform, where everyone would need to connect to your platform before they can participate in a data flow. This is likely to fail because it doesn't allow for the expected "network effect" to take hold, and systems that are decentralised at their core would then quickly win out against any centralised solution.

The point is not to build a platform, but to build an ecosystem. And to build an ecosystem, an open mindset is required.

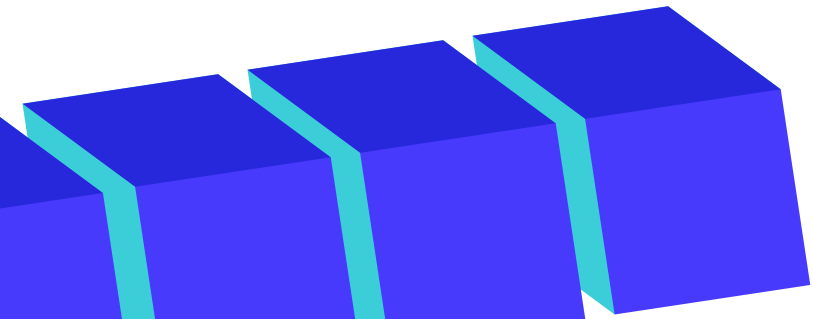
Contrary to many other use cases where blockchain is involved, very little data will be stored on-chain. This is because decentralised systems only use the blockchain as a mechanism to securely distribute the public keys and certain other metadata attributes. At no point in the flow of data does a decentralised identity solution expect private data to be stored on a blockchain. Such private data should always flow peer-to-peer between the involved parties only.

Also, because of this, the amount of data published to the blockchain remains limited and rather static. Rarely does a public key change, and this allows us to consider higher volume use cases than those often supported by existing blockchain solutions. Again, the vast majority of the data flows peer-to-peer between the different parties, not on the blockchain itself.

It also means that any shared blockchain infrastructure would benefit from being public by default, so it is unlikely that a private blockchain network is needed. This helps with getting the opportunity off the ground, as there's less infrastructure to worry about. Often, existing public permissionless networks can be used.

All of this helps us focus the use cases away from platform and infrastructure thinking, and towards that which immediately brings value to the table. If you have data of value or know of data that would be valuable to you, it shouldn't be too difficult to identify who's involved in such a flow. Once you've identified that, the next step would be to reach out to see if there's a shared interest in using decentralised identity solutions to better connect and share this data.

The argument is that decentralised identity solutions make it quicker, easier, and more efficient to manage and share data with fewer lock-in risks, fewer dependencies, and better compliance. Decentralised identity will unlock opportunities that previously came with too much overhead, manual intervention, and unclear provenance.



Conclusion

Through this ebook we've described limitations and risks associated with existing identity solutions, and we've argued why decentralised identity helps address these. While these new solutions clearly can offer a better experience for the end user, the important thing to remember is that the opportunity is much larger than just a better sign-up and sign-in flow.

Perhaps, in a few years, we will look back and attribute only a small fraction of the value unlocked by these new systems to those involving direct human interaction. The potential for devices, organisations, and business processes to directly establish trusted connections with the outside world is immense. As highlighted in the introduction, we expected this to grow into a multi-trillion dollar opportunity by 2030, with aggressive year-over-year growth.

The approach should be based on understanding that to succeed with decentralised systems, you must think in terms of ecosystems and not platforms. The focus should be on unlocking data opportunities, either by releasing data held in internal data silos or by helping to standardise how such data can securely flow between participants.

Working with Web3 Labs

Customers

vodafone

OSTTRA

J.P.Morgan

Palm NFT Studio

r3.

Partners

Microsoft

vmware®

Infosys

ENTERPRISE ETHEREUM ALLIANCE

CONSENSYS

parity

www.web3labs.com