Enabling Trust in Data Exchange for Materials 4.0

Commissioned by the Henry Royce Institute for advanced materials research and innovation

July 2021

In collaboration with

HENRY ROYCE INSTITUTE

CATAPULT Digital

About roadmapping and landscaping

This report is commissioned by the Henry Royce Institute for advanced materials as part of its role around convening and supporting the UK advanced materials community to help promote and develop new research activity.

The overriding objective is to bring together the advanced materials community to discuss, analyse and assimilate opportunities for emerging materials research for economic and societal benefit. Such research is ultimately linked to both national and global drivers, namely the transition to zero carbon, sustainable manufacturing, digital and communications, the circular economy and health and wellbeing.

Contents

Foreword

Materials 4.0 aims to radically change the rate and responsiveness of materials innovation, increasing the impact it has on society and the economy.

The materials and manufacturing sector forms 15% of UK GDP and has a key role to play in achieving technological and societal goals such as the transition to net zero carbon. Such advances will require the approval and acceptance of materials that are either at the early stage of development or yet to be discovered.

One of the primary challenges to the rate of development for new materials is the slow and complex systems that are required to share sensitive data between parties. These obstructive systems provide no insight into the technical suitability of these materials, yet can often lead to delays or even failure due to an inability to agree the terms for sharing of key data.

The aim of this paper is to guide the thought processes of the materials community on methods for addressing this barrier. This will be achieved by providing expert opinion on the existing and emerging technologies in digital security and data trust, for applications in materials and manufacturing.

Professor Iain Todd

Project Champion and Scientific Lead for Materials 4.0 roadmap, Henry Royce Institute

Executive summary

In materials development, there's always a reason not to exchange data. This paper explores how distributed systems can remove some of the barriers and where this has been done in other sectors.

When collaborating in a decentralised fashion, parties must be able to trust one another in the absence of a traditional central authority to govern interactions.

A mutual set of rules and policies must be established to govern processes, which are then agreed by any participating entity. This is typically what is meant by consensus at a human level: a mutual agreement between parties.

Distributed systems have recently come into focus due to activity in cryptocurrencies such as Bitcoin or Ethereum. However, there are several other distributed systems tools that can be used to align the state-of-knowledge of unknowing or untrusting parties. These are immediately applicable to the challenges surrounding Materials 4.0, and should be explored in an appropriate fashion.

Individual and corporate-level curiosity has already translated into a genuine interest into the wider applicability of some of the core technologies underpinning these inventions – most notably the blockchain, a form of distributed ledger.

Distributed systems are now having an impact in diverging industries, such as pharmaceuticals, construction, nuclear, aerospace and food and drink. For the materials engineering and manufacturing sector, there is a distinct opportunity to innovate and explore how multiple parties using distributed systems can improve trust across supply chains.

Challenges of trust and data exchange in materials

1

ACCELERATE THE DISCOVERY, INNOVATION AND VALIDATION OF NEW MATERIALS.

Challenges of trust and data exchange in materials

Materials 4.0 proposes a digital materials revolution to accelerate the discovery, innovation and validation of new materials.

Its goal is to maximise the value of materials data and link the digital and physical via cyber-physical systems for prediction, classification and control of material performance. This interdisciplinary approach will support improved products with longer lifetimes and enable their intelligent reuse, bridging the current gap between opportunity and capability.

Through the Materials 4.0 programme, organisations will become data-centric, but within materials science and engineering culture, there are frequent reasons not to share data – fundamentally due to trust.

Materials are essential to product performance, safety, lifespan and sustainability. As a result, organisations that develop new materials are very protective over their data. However, materials innovation is slow, with development cycles presently measured in decades, rather than months or years.

Therefore, there is a huge opportunity in creating trust and data exchange. Not just to reduce the development lifecycle, but also improve material performance, reduce the cost of materials development, reduce the environmental impact of materials and improve public trust in materials.

HOW DOES CONSENSUS ENABLE TRUST?

When collaborating in a decentralised fashion, with the absence of a traditional central authority to govern interactions, parties must be able to trust one another. Instead, a mutual set of rules and policies are established which govern processes, which are agreed by any participating entity. This is referred to as consensus, a mutual agreement between parties that is enacted at human or organisational level.

For example, this could involve a consortium of parties using steering groups to agree on a new standard for material quality for a particular industry application or the United Nations Security Council agreeing on appropriate sanctions for a country that has been acting in bad faith.

On the other hand, distributed ledgers use consensus algorithms. These are the mathematical processes used by nodes in a distributed system to ensure each one holds an identical replica of the same data, ensuring the network is operating as a whole. These computer-level systems and processes strongly support and enhance human level agreement and collaboration.

A distributed ledger is immutable and provides an indisputable audit trail of transactions and entries within a system. While this doesn't guarantee that all information is correct, it does provide a deterrent for bad actors since the record of a potentially malicious event cannot be removed. Distributed systems can also incorporate verifiable credentials to validate the identities of actors. This provides an additional layer of security without necessarily needing an outside authority to perform these checks.

The multi-party nature of these systems facilitates collaboration and helps build greater trust over time. Interacting parties share a common goal of wanting to keep their data secure and ensure the integrity of the records, so have a shared incentive to promote good behaviour while limiting any attempts to undermine this. Using appropriate technologies, such as distributed ledgers, identities or file systems inherently helps to build trust and credibility between parties within the ecosystem.

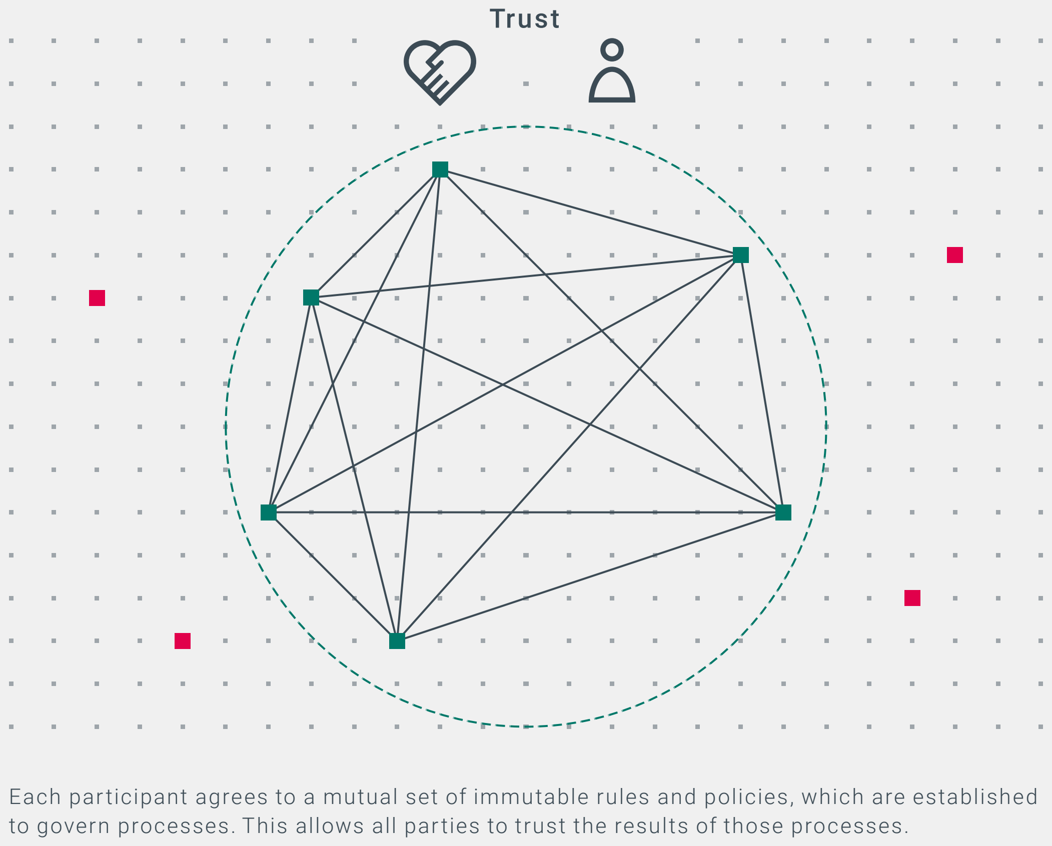

Trust
Each participant agrees to a mutual set of immutable rules and policies, which are established to govern processes. This allows all parties to trust the results of those processes.

Figure 1 - How does consensus enable trust?

RESPONSIBLE DIGITAL INNOVATION

According to estimates by the Cambridge Centre for Alternative Finance, the global electricity consumption of Bitcoin mining is now greater than the total amount of electricity consumed in Argentina or the United Arab Emirates. This high energy usage results from using proof-of-work (PoW) algorithms to ensure the network operates with a singular view of truth (consensus). Proof-of-work is a mathematical means by which one party proves to others that a certain amount of computational effort has been expended to secure the system. As a result, users are rewarded for expending large amounts of computation, which uses a large amount of energy. This problem is compounded in countries with non-green energy mixes, as well as those where energy costs are subsidised (e.g. Venezuela, Iran and China) so that market behaviours no longer limit uneconomic activities.

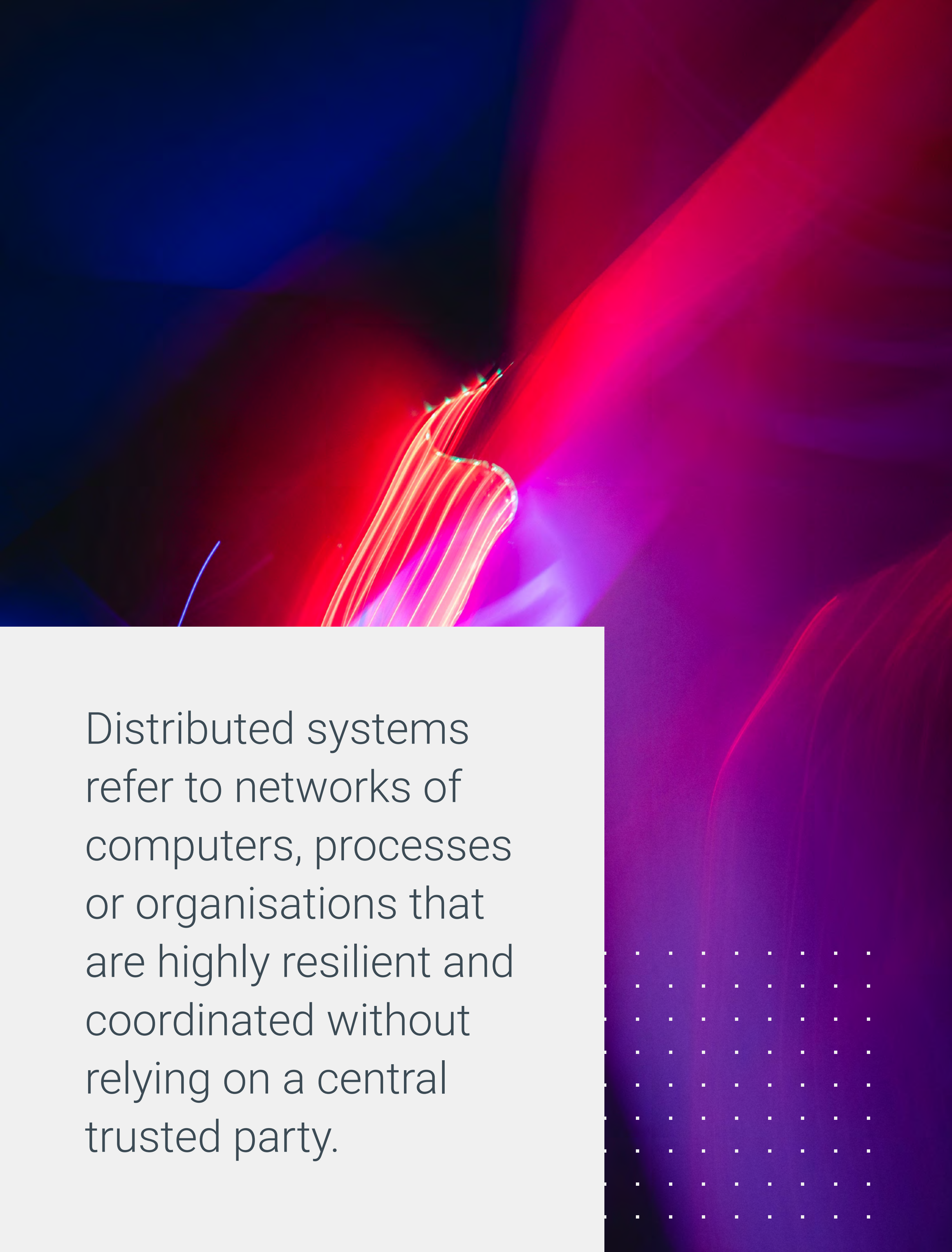
Alternative, energy efficient consensus algorithms are an area of active research and development within the global distributed ledger community. In fact, the original consensus algorithms dating back to the late 1980s consumed minimal energy. Although these algorithms have evolved, they are still in use today by many projects working in the private ledger or enterprise ledger space.

Public distributed ledgers/blockchains are also moving away from proof-of-work. Instead, well-known projects such as Ethereum 2.0 are transitioning to proof-of-stake algorithms, where users put value at risk instead of expending computational power. However, it is unlikely that Bitcoin will move away from PoW mechanisms, due to internal political decisions within the project.

Distributed systems play a significant role in ensuring that the right parties have access to the right data at the right time. This includes verifying assertions and data to prove that products have been manufactured in a way that results in a positive environmental impact. It is important to ensure that the use of digital technologies does not increase the carbon footprint of an existing operation.

As a result, when designing a distributed system, it is important to:

- Ensure that data is stored in an environmentally positive (physical) location – this could be in a data centre that utilises 100% renewable energy
- Treat all data as you would a physical asset: understand the economic and environmental costs of storage, store only what data is necessary and discard the excess
- Understand that distributed ledgers are only one tool within the distributed system toolkit and must be used appropriately alongside others
- If a ledger forms a necessary part of the final infrastructure, select a consensus algorithm that is appropriate for the intended network topology and participant mix



Distributed systems refer to networks of computers, processes or organisations that are highly resilient and coordinated without relying on a central trusted party.

INTRODUCTION TO DISTRIBUTED SYSTEMS

The terminology and concepts behind distributed systems date back to the earliest days of networking in the 1960s. Then Paul Baran of RAND Corporation was exploring topologies capable of withstanding a nuclear attack and invented the concept of packet switching[1]. This concept was simultaneously conceived and named by Donald Davies at the UK's National Physical Laboratory, whose work fed back into the creation of the ARPANET – the precursor of today's internet.

At their most basic level, networks consist of interlinked nodes, where nodes are points of message origination, consumption or redirection. At a physical level, nodes usually refer to individual servers[2], but these can be thought of as representing individual stakeholders in the network, such as a company, laboratory or regulatory organisation.

The arrangement of these nodes and interlinks can take the form of three different possible topologies, as described by Baran:

- **Centralised** - a typical client-server infrastructure where all messages are relayed by a central coordinating point, providing a single point of high vulnerability
- **Decentralised** - a hub-and-spoke model, wherein peripheral nodes connect to central hubs forming small clusters, which then interconnect. This is the shape of today's internet, where targeted attacks on major hubs can prevent more distant nodes from communicating
- **Distributed** - a totally flat topology, without any hierarchy and far more interlinks. Messages can route around failing nodes through redundant connections

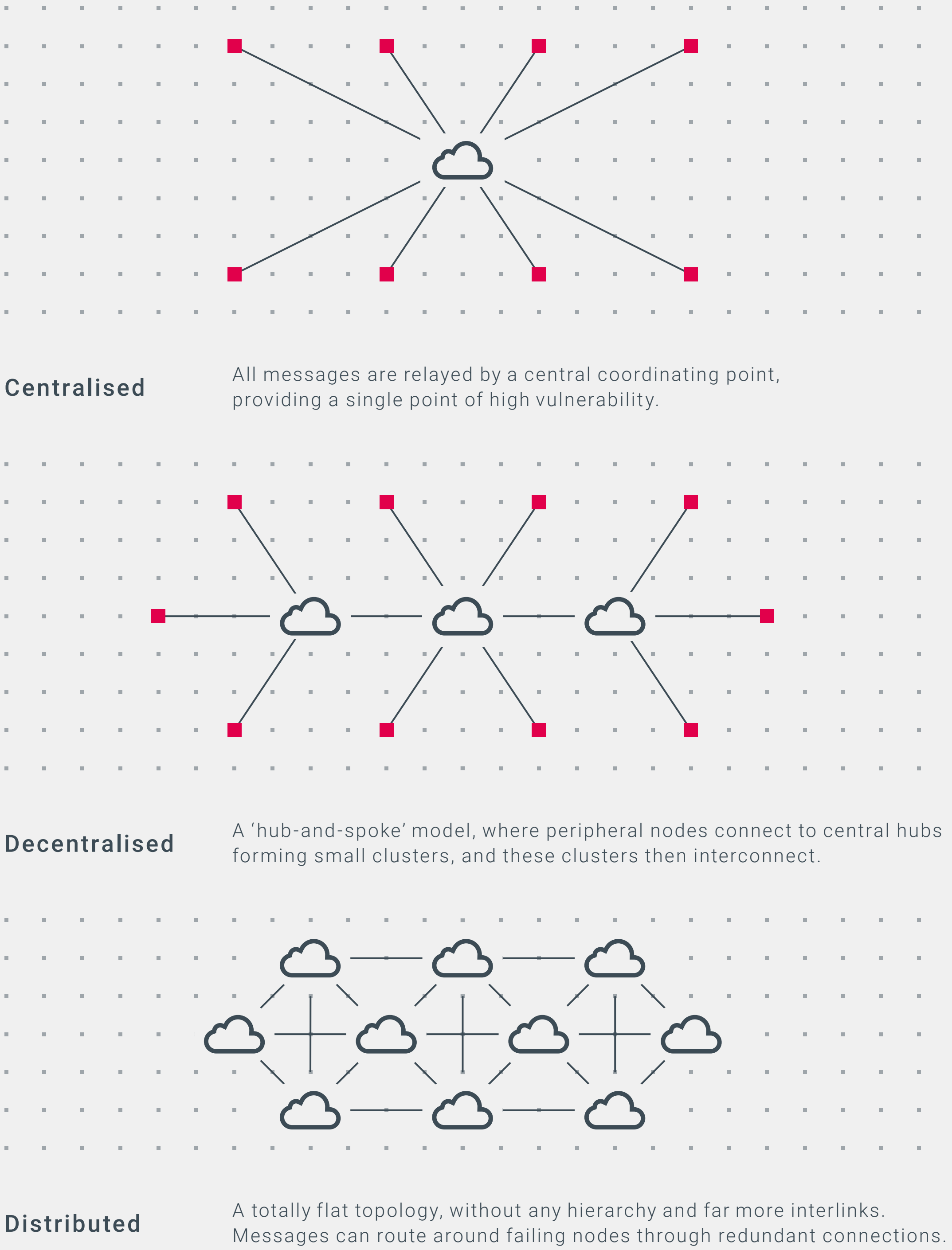
Distributed systems have recently come to the fore due to activity in cryptocurrencies such as Bitcoin or Ethereum. Individual and corporate-level curiosity or speculation has translated into a genuine interest into the wider applicability of some of the core technologies underpinning these inventions – most notably the blockchain, a form of distributed ledger[3].

While distributed ledgers are a fantastic invention in computer science for aligning the state-of-knowledge of unknowing or untrusting parties, they are not the only offering from this older diverse and ever-growing field.

Other distributed systems tools include:

- Trustless file-sharing
- Process automation to align independent parties
- Decentralised asset registries with embedded provenance
- Self-sovereign identity to return privacy and control back to stakeholders
- Privacy-enhancing cryptography to extract value from raw data without revealing the underlying data

A number of these are immediately applicable to the challenges of Materials 4.0, and should be explored accordingly.



Centralised

All messages are relayed by a central coordinating point, providing a single point of high vulnerability.

Decentralised

A ‘hub-and-spoke’ model, where peripheral nodes connect to central hubs forming small clusters, and these clusters then interconnect.

Distributed

A totally flat topology, without any hierarchy and far more interlinks. Messages can route around failing nodes through redundant connections.

Figure 2 - Introduction to distributed system

The potential benefits and impact to industry

2

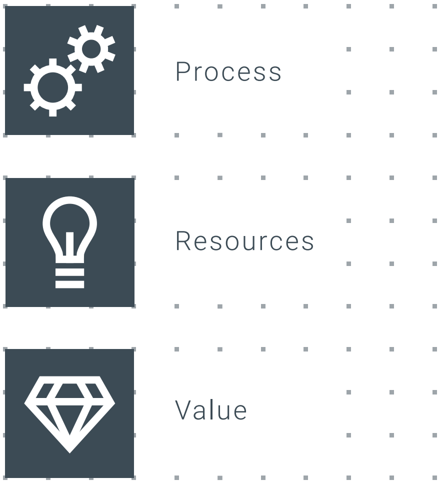
The potential benefits and impact to industry

There are three main benefit categories of trust and data exchange that impact the materials industries: **operational, strategic and consortium.**

OPERATIONAL BENEFITS

Operational benefits can be separated into three different aspects:

- **Optimisation of processes** – data flow within a node (factory, facility, organisation etc.). By driving trust and data exchange within a node, greater insight is created within production systems, helping to reduce cost, enhance quality and improve delivery schedules. Additional data insight creates benefits for audit and compliance, while checks could be done remotely and with more frequency to provide additional value.



- **Optimisation of resources** – data flow across a network. By driving trust and data exchange across a network, human capital and material resources are optimised for greater productivity and reduced waste. Overall, this has a positive impact on sustainability, particularly around greenhouse gas emissions and should be a strategic priority for UK industry. By defining provenance, industry would receive full, auditable maps of supply networks (excluding identifiers), enabling accountability and an overview of resilience.

- **Creation of new value** – data insight. Users access faster, cheaper finance through data insight which can be used to incentivise delivery of sustainable development goals. The additional insight in audit and compliance provides a more accurate risk profile, which could also reduce the cost of insurance. All this adds up to a combination of reduced cost and higher profitability, as well as the potential for revenue growth.

STRATEGIC BENEFITS

Strategic benefits are longer term and less measurable, but have greater potential impact for industry. These include:

- Data driven policy development and faster delivery into industry
- Scenario planning and modelling
- Risk management
- Increasing public trust in new materials development
- Accelerating the net zero agenda

Society

Net zero

Trust

ECOSYSTEM BENEFITS

These are achieved when multiple parties find effective ways to develop trust and exchange data:

- **Access to new markets** - through developing domain knowledge in different materials/ sectors and removing the barriers to entry
- **Creation of new services** - developing new materials and associated shared data into a higher value, higher impact offering, testing the business model options before going to market and having faster product development cycles

New markets

Services

Business model

INFORMATION FLOWS AND TECHNOLOGIES

There are a number of information flows required to achieve the intended goals of Materials 4.0 Information can pass through authorised central bodies imbued with the regulatory or industry power to ensure total data capture and standardisation. Alternatively, they can pass directly between those who have good reason to access the data, as and when required. This disintermediation and direct interconnection is also part of the distributed systems proposition.

We will now explore three of these technologies with accompanying thought exercises on their applicability to Materials 4.0.

Distributed ledgers are time-sequenced, non-repudiable, and unforgeable records of state. These can be combined with programmable logic to regulate state transitions and cryptographic proof of existence and fed by either human or machine-to-machine information. These features can, for the first time, minimise knowledge imbalance and coordination costs in the digital space.

Appropriate design and deployment of distributed ledger technology (DLT) can enable multiple independent parties to come together and track events around material lifecycles, with added verification of facts to enhance trust. A number of features at the heart of these systems are critical to enabling this future:

- The consensus process of the ledger ensures all participants share a non-conflicting, singular view of the history and current state of the system
- Each record is individually signed with unforgeable cryptographic signatures, uniquely and irrevocably associating every record with its signatory or signatories
- Cryptographic methods can be used to keep certain data private or to prove the existence of critical information at a given moment in time

With verifiable and trackable knowledge around the production, development or usage of novel materials, it will be easier to hold parties responsible for their activities and harder to fall out of compliance. The trackable and controllable movement of data assets can also create a fair and balanced knowledge (or financial) economy for participants in the Materials 4.0 ecosystem.

Distributed file systems are some of the older and more familiar technologies in this field. These enable parties to locate and share larger digital objects as seamlessly as accessing a file on a local machine. Benefits include the ability to share storage resources between stakeholders to increase availability and resist single points of failure that may result in data loss.

Implementing these tools requires a balance between security, resilience, performance and multi-stakeholder concurrency (for example, handling the ability for one person to update a file while another is actively reviewing it). Fortunately, there are multiple offerings available in the field, each striking their own balance.

There are various distributed file storage models in use today:

- **Networked storage** is perhaps the earliest model of multi-party file access and dates back to the 1970s. This model of file storage has multiple client systems attached to central servers, with the ability to treat the files as if they were stored locally. Familiar versions of this approach include the network file system (NFS) from Sun Microsystems, Apple file protocol (AFP), and server message block (SMB) from IBM (modified by Microsoft for Windows systems).

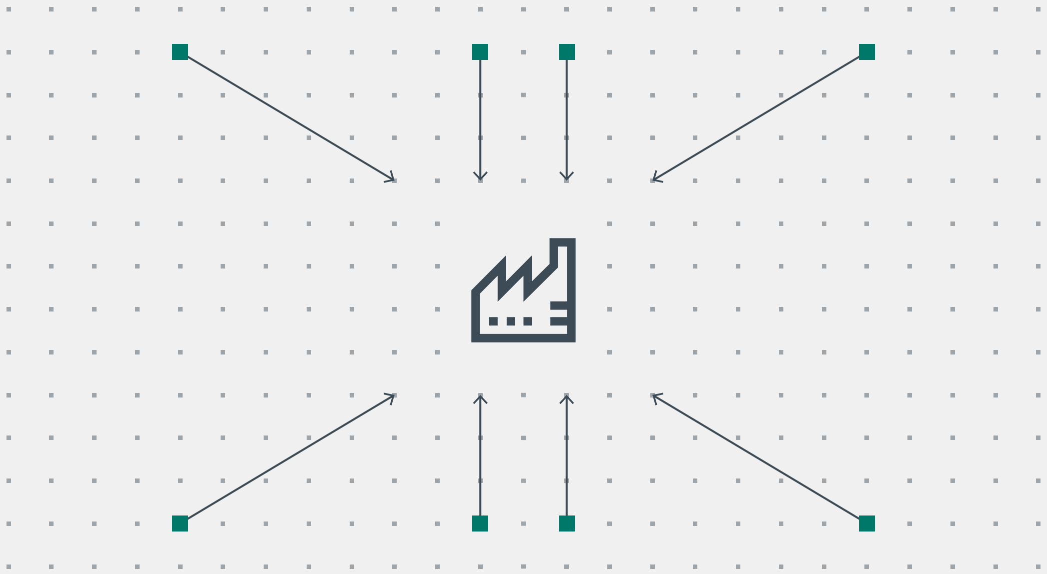
  This type of storage is typically deployed within the bounds of a single organisation and managed by a system administrator, who can modify user privileges for files or directories. This is most similar to the centralised network model described by Paul Baran.

- **Cloud storage** is an evolution of the networked storage model, provided or managed by a trusted third party (Amazon, Google, Microsoft, DropBox, etc.). Through user-friendly interfaces, data can be sent and retrieved in near-seamless manner via proprietary data centres. Within these data centres, the data is then striped or distributed across multiple servers for load balancing and redundancy. Cloud sits between the centralised and decentralised network model where all trust is placed in the behaviour of a third party, but data is split across hubs and multiple independent parties can access the data with the correct permissions.

- **Peer-to-peer storage** seeks to reflect the fully distributed network model, with each node having access to each data store (with appropriate permissions) without traversing an intermediary third party. This model was famously deployed in the guise of BitTorrent for illegally sharing music and Hollywood movies in the late 1990s/early 2000s. However, this technology did demonstrate the power and resilience of this new approach: segmenting (chunking) files into multiple blocks, individually addressing these with hash-based algorithms, and distributing look-up tables of blocks that constitute a final file. In this way, files can be simultaneously retrieved and re-transmitted by multiple peers until the full inventory is achieved, reconstituting the final file. Modern evolutions of the BitTorrent protocol include the InterPlanetary File System (IPFS) and DAT, which have far more mainstream uses.

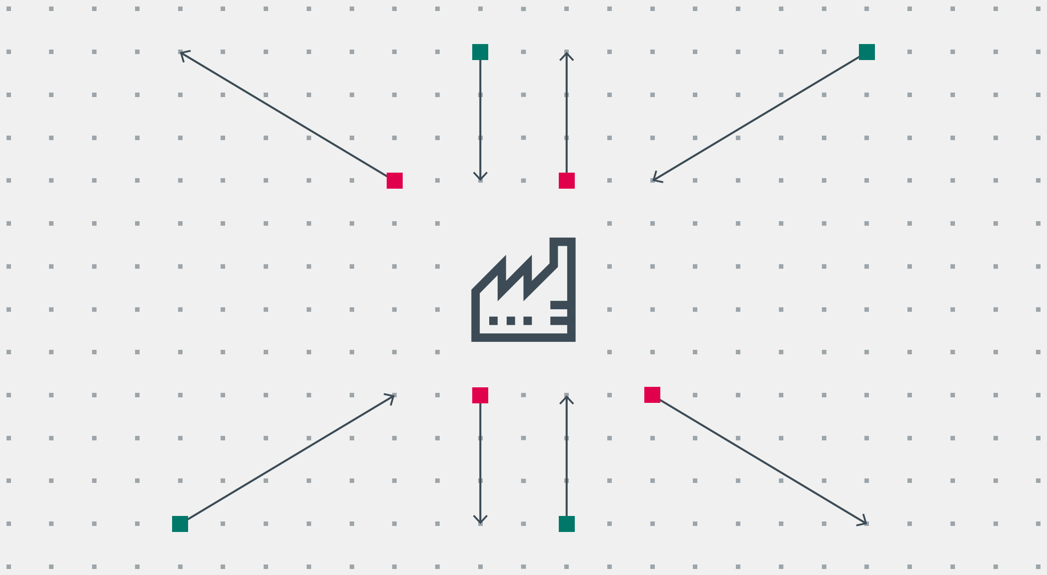
Peer-to-peer storage has the most to offer those working in Materials 4.0, as it closely reflects the real-world organisation and interaction of the diverse stakeholders in this field.

Privacy-preserving cryptography involves advanced mathematical methods for enabling untrusting parties to gain new insights from their shared data. Traditional thinking would have both sides transmit their data to a central location where analysis can be performed or to directly exchange raw data between each other. This is, of course, fraught with issues around trust, copying, leakage or loss, which require lengthy negotiations and legal contracting. This frequently means that data is never shared. Instead the appropriate application of new cryptographic techniques can enable data processing while also preserving data privacy, providing net benefit for all concerned. This can ensure the Materials 4.0 ecosystem realises a whole that is greater than the sum of its parts.



Coordinated attack (Victory)

In a distributed system, disparate stakeholders can be digitally coordinated around a common goal, with each able to trust the contributions of the other.



UnCoordinated attack (Defeat)

Without the ability to properly coordinate or trust each stakeholder's input, it becomes far more difficult to achieve consensus.

Figure 3 - Information flows and technologies (Byzantine generals)

Examples include:

- **Federated learning (FL)** iteratively trains a machine learning algorithm on data held locally so that only the model is transmitted between parties, rather than the underlying raw data. This iterative approach can leak some information about the data, but can, for example, be used to train models from thousands of edge devices.

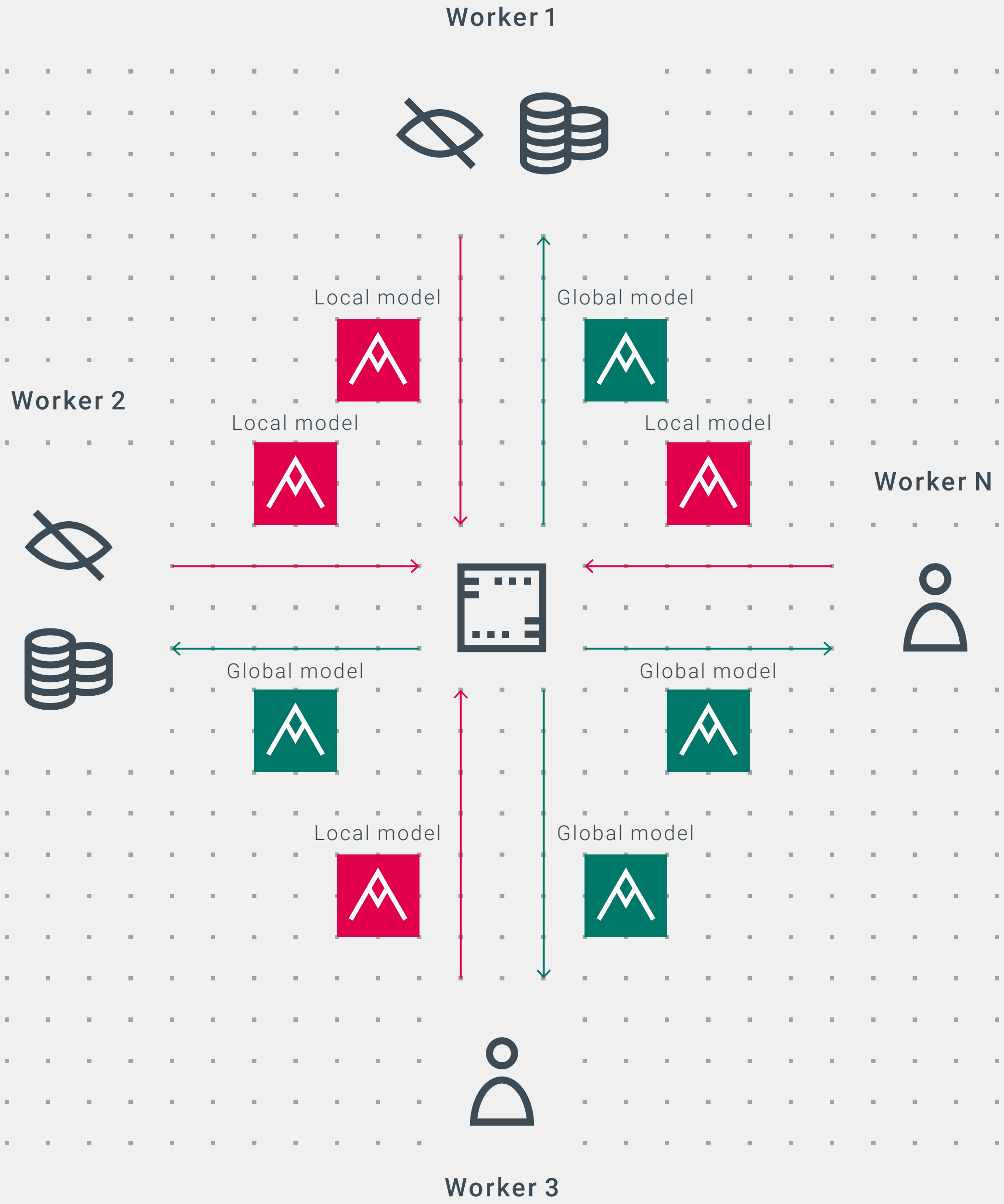


Figure 4 - Federate learning (FL)

- **Multiparty computation (MPC)** is a more secure version of the federated learning concept. MPC does require more computational overhead but can produce outputs without revealing any intermediate steps.

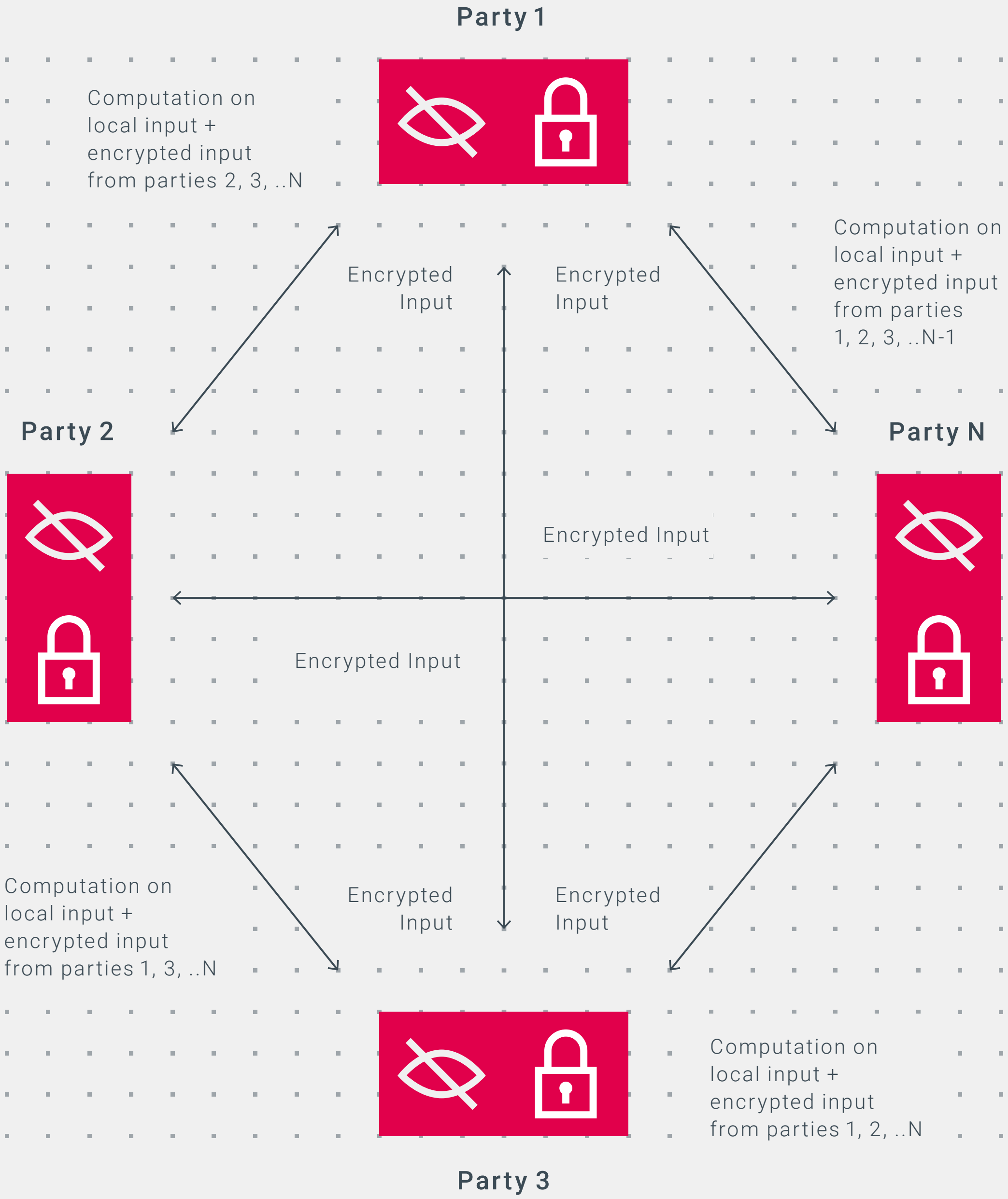

Party 1
Computation on local input + encrypted input from parties 2, 3, ..N
Encrypted Input
Encrypted Input
Computation on local input + encrypted input from parties 1, 2, 3, ..N-1
Party 2
Party N
Encrypted Input
Encrypted Input
Computation on local input + encrypted input from parties 1, 3, ..N
Encrypted Input
Encrypted Input
Computation on local input + encrypted input from parties 1, 2, ..N
Party 3

Figure 5 - Multiparty computation (MPC)

- **Homomorphic encryption (HE)** enables calculations to be performed in a locked vault on encrypted data without performing decryption at any stage. This is the most computationally intensive technique, but can be used to extract great value from secret datasets that have been prepared appropriately – for example benchmarking performance data between different organisations.

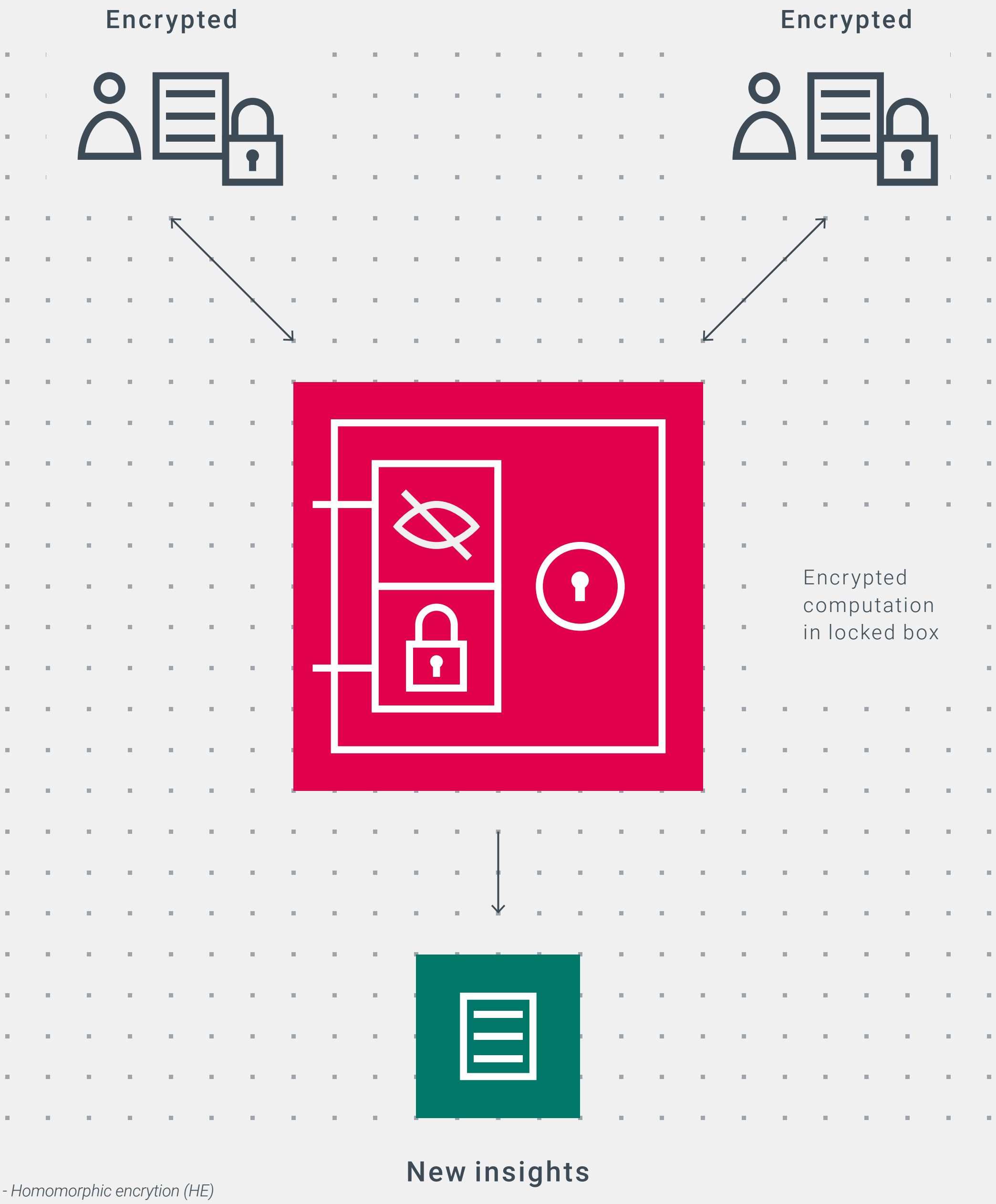


Figure 6 - Homomorphic encrytion (HE)

- **Zero-knowledge proofs (ZKP)** allow one to prove an assertion about data without revealing the underlying data itself – for example: mathematical proof that a machine followed the correct manufacturing process.

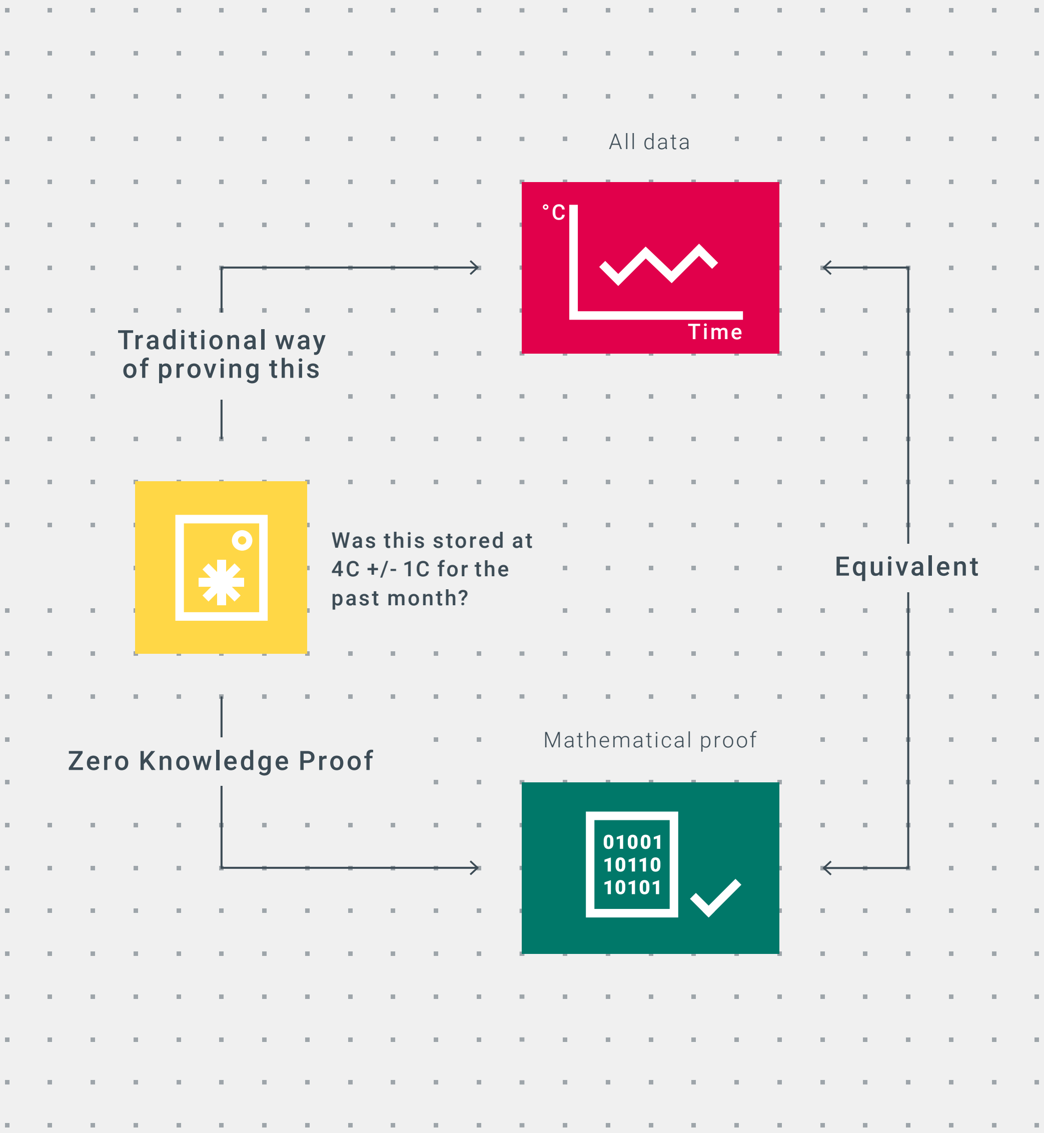
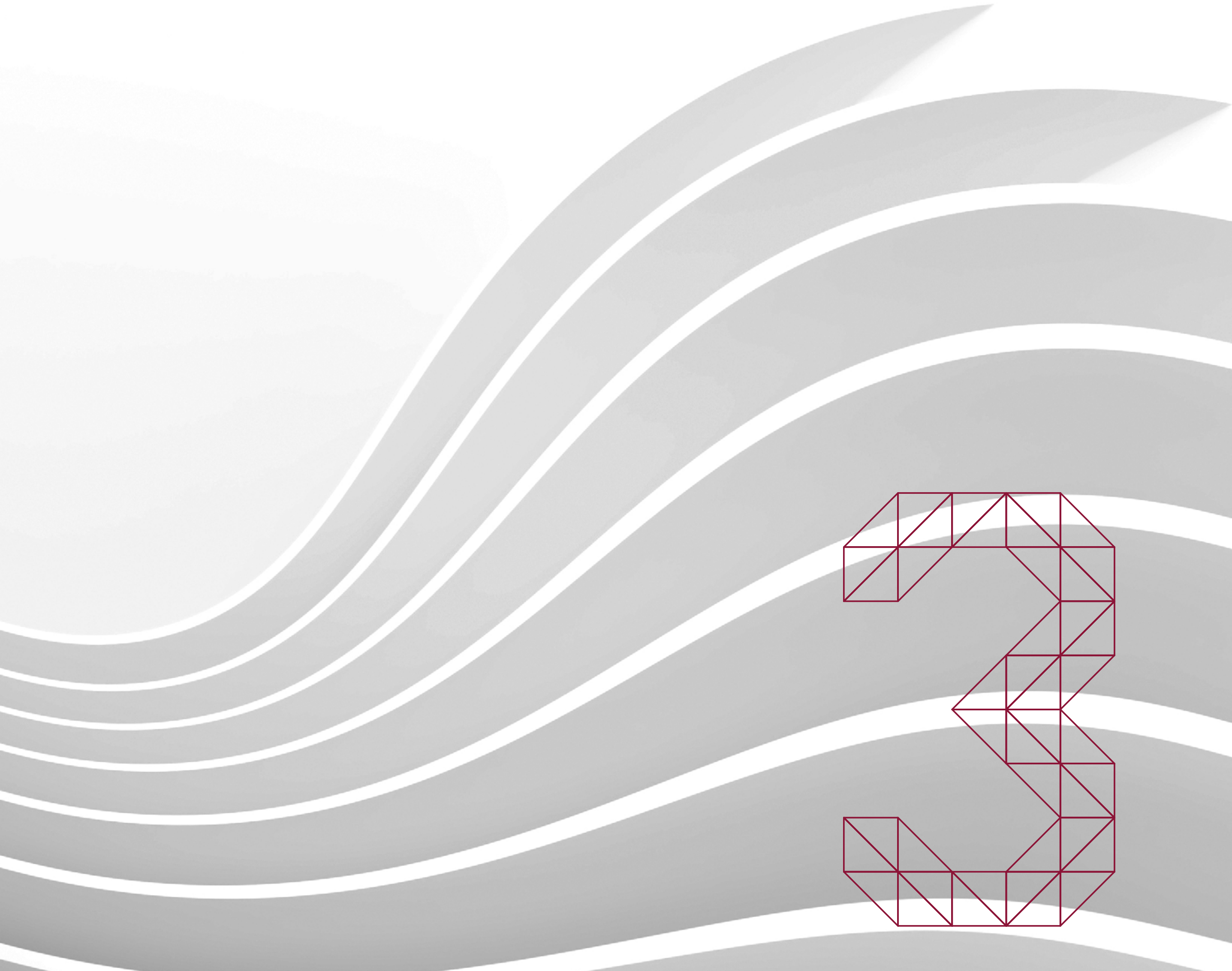

All data
°C
Time
Traditional way of proving this
Was this stored at 4C +/- 1C for the past month?
Equivalent
Zero Knowledge Proof
Mathematical proof
01001
10110
10101

Figure 7 - Zero-knowledge proofs (ZKP)

Key enablers to create impact from distributed systems



Key enablers to create impact from distributed systems

MULTI-TECHNOLOGY APPROACH

Distributed systems do not exist in a vacuum. These technologies are extremely good at what they do, but they need to sit at the right layer of the tech stack to provide maximum value. This is at an infrastructural layer between the data sources (humans or machines for example: Internet-of-Things (IoT) devices) and the data analysis or presentation layer.

In this position a distributed ledger can reinforce facts about data for example: what happened and when, and who or what was involved, as well as ensuring data consistency and formatting standards. Whether this data comes from a cloud source, a device or an on-premises server, a peer-to-peer distributed file system can ensure accessibility, redundancy and uniqueness of representation.

With trusted, consistently formatted, available data and the appropriate privacy-preserving encryption, a variety of machine learning or AI techniques can be applied. These can be used to optimise processes and extract insights, guiding both policy and individual business decisions.

GOVERNANCE

While distributed systems can provide the necessary infrastructure required to digitally coordinate disparate stakeholders around a common goal, the human-level coordination of these parties still typically requires mutual agreement on a suitable legal and organisational framework to achieve success.

The term governance describes the process of organising and managing the collective behaviours of groups of persons or organisations. Creating a suitable governance framework for deploying and managing shared digital infrastructure between disparate stakeholders can be fraught with high costs, lengthy debate and organisational frictions.

Digital Catapult has undertaken exploratory work in the field of developing appropriate frameworks for businesses using distributed systems. This has drawn upon the expertise of numerous stakeholders in the field, from academics and legal professionals, to innovators and traditional industry.

The result is a collection of core topics and prompts that guide groups to develop an initial charter, documenting the best collective understanding of the issues at hand. This serves as a starting point for further discussions and amendments during the execution of experiments with these technologies, prior to engaging legal experts for final comprehensive contracts.

Translating these human-level rules into the automated machine-level rules governing data exchange, permissioning and onboarding or exclusion of participants, is a further challenge to active research and development.

BUSINESS MODEL AND INCENTIVISATION

The challenges of developing suitable human-level governance models include adequate accounting for each stakeholder's voice, role and concerns. Similarly, any successful distributed system must also present suitable incentives and opportunities for participants to benefit economically, socially or otherwise.

For the first time, distributed ledger technology allows disparate and untrusting stakeholders to agree on common resource existence, ownership and rules of exchange. All while keeping the coordination costs low compared to traditional paper or legal methods.

One way distributed ledgers reduce coordination costs is by integrating mechanisms for tracking and assigning value to data flows, just as those seen in Bitcoin, Ethereum and many other blockchain projects. However, frictions can arise at the interface between the internal economy denominated in a virtual cryptocurrency, and the external economy denominated in national fiat currencies. There are solutions to this challenge, but it must be considered before considering payment mechanisms within the infrastructure itself, beyond the data monitoring and exchange functionality.

Distributed infrastructures can facilitate a wide range of market structures by combining distributed ledgers that track the state of verifiable assertions with file systems that enhance data availability and privacy-preserving computation. This economic layer must of course interact with the governance layer, but the problems of our current economic systems (being fairness, transparency and efficiency) could be largely diminished through the appropriate application of distributed technologies.

The goal of any well-balanced distributed infrastructure should be to enable novel non-zero-sum economics, or so-called co-opetive[4] business models. This is where competing stakeholders accept that pursuing traditional market capture and monopoly is less efficient than cooperation in certain economic areas. This ideology should be considered when modelling and implementing a fair data economy for the Materials 4.0 ecosystem, based upon the trackable, controllable and accountable exchange of data assets between participants.

INTEROPERABILITY

This is the ability for technically different systems to interact, communicate and exchange or operate on data regardless of the provenance of the system or who manages it.

Interoperability is usually enabled by the standardisation of data formats and communication protocols.

Common examples include:

- HTML for presenting web pages
- JSON and XML for data mark-up
- Emoji libraries, ensuring that a thumbs up looks the same to everyone around the world
- Email relay protocols that ensure delivery, regardless of the provider or device
- The TCP/IP protocols that underpin the internet

These can either emerge through consultative work and open publication (Open Standards), or through first-mover advantage and market dominance, forcing others to modify their tooling to ensure compatibility, if not interoperability. The latter explains the occasionally jarring formatting errors should the same *.ppt presentation file be opened in the standard leader Microsoft PowerPoint versus compatible software such as Google Slides, Apple Keynote or OpenOffice Impress.

The world of distributed systems contains examples of both types of standardisation leading to interoperability, with one large difference: most of the software tooling is entirely open source. Most open source software development also takes place in public, and significant changes in behaviours (breaking changes) are often signalled well in advance so that others who rely on the software can be ready.

Where open source denotes that the underlying code for a particular software module or functionality is available for others to review, an open API describes the open documentation available for an application programming interface. Even closed-source, or proprietary software such as Microsoft Windows, Facebook or Google Maps has documented and public-facing APIs to enable developers to write applications that communicate with the underlying software. Any software application downloaded on an Apple iPhone takes advantage of its open APIs to create novel and useful functionality.

The open nature of most software development and communication practices within the distributed systems community enables a potentially high degree of interoperability and can future-proof systems built today.

PERMISSIONING

The final enabler to create impact from distributed systems is the ability to limit permission data access and visibility to specified parties. These controls may be difficult to achieve with the current state of the art, but depend upon the levels of visibility acceptable to participants.

For example, if transactions are being trustlessly recorded and enabled by a distributed ledger shared by a number of parties, it may not be possible to blind certain participants to the existence of a transaction without breaking consensus about the state of events.

Potential solutions come in the form of three different approaches: network topology, advanced cryptography and permissioning machines.

Topological approaches have been offered and developed by the wider distributed systems community and form the basis of significant projects. These rely upon purposeful fragmentation of the ecosystem to create private bidirectional channels between parties, or smaller working groups of organisations that can choose to interact with the larger external ecosystem, but whose communications are otherwise invisible to them.

These smaller groups achieve consensus and data distribution within their group to minimise knowledge imbalance and enhance fault tolerance, but if they wish to communicate with another group, they must route through an overarching interoperability layer. The technical mechanisms are complex, but already in active development with terminology such as shards or parachains.

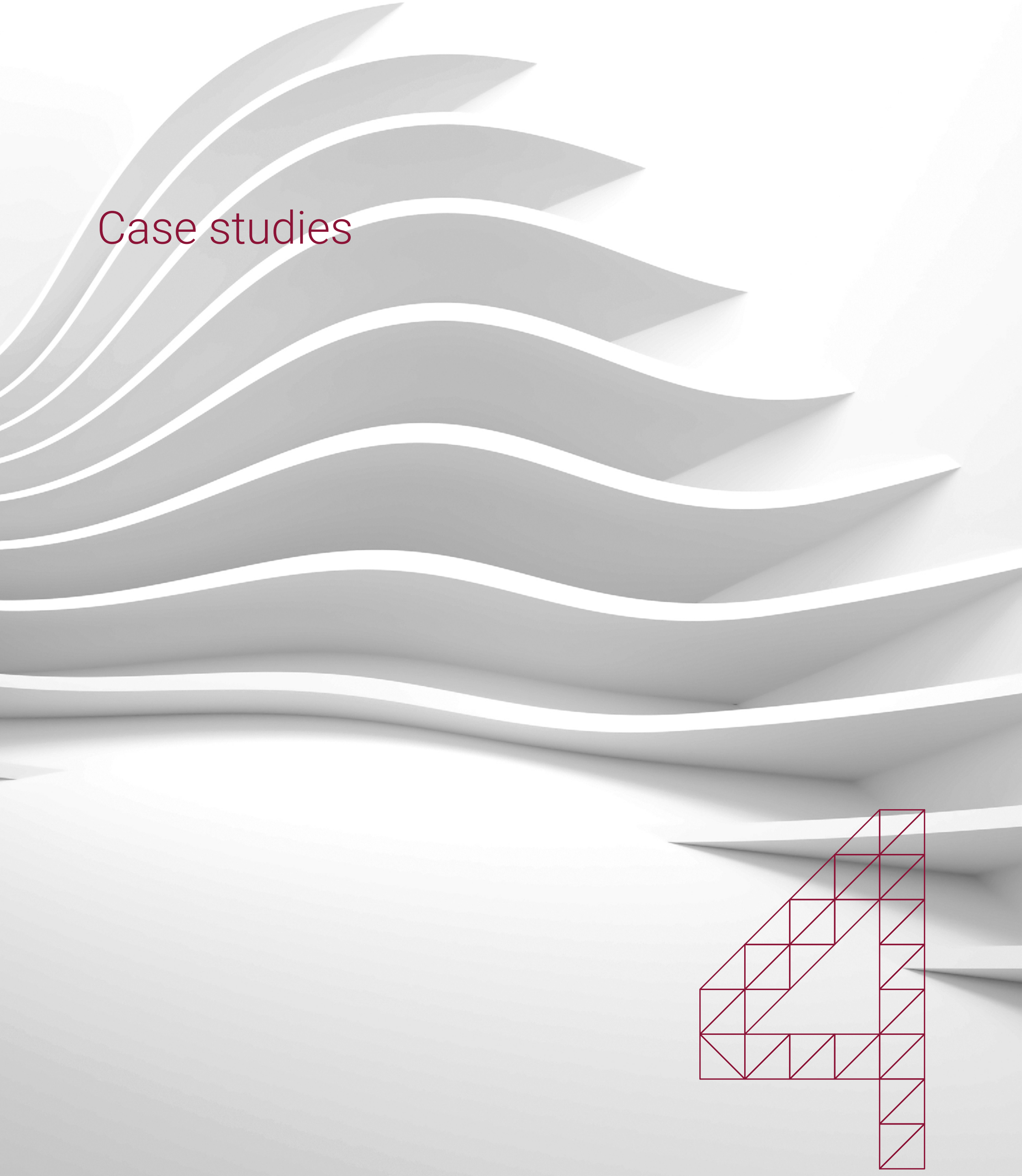
In addition to topological design considerations to enable permissioning, advanced cryptographic methods can be employed. As long as the existence of messages does not have to be kept secret, cryptography can ensure that only the recipient of a data payload is capable of reading it. Alternatively, data can only be processed using the methods described previously, such as homomorphic encryption or multi-party computation.

Finally, autonomous software can be deployed across the infrastructure that enforces permissioning in much the same way that access control is implemented in centralised systems such as Sharepoint or DropBox. Using a suitable permissioning markup schema such as XACML and distributed identifiers for organisations or individuals, the system ensures that only the appropriate parties can access certain data without a single central controller.

Data access and permissioning within a distributed system that employs distributed ledgers to track attribution and interaction have the added security of persistent surveillance. This approach fits alongside hiding the data with private channels and locking the data away with cryptography. Should anyone overcome these two aspects of security, the ledger will still be present to record liability for actions, so that the governance layer can seek recompense. In some cases, misdeeds can also be handled automatically by governance rules built into the system, such as by sanction or ejection to ensure that the disincentives for misbehaviour are balanced against potential benefits.

In this way, well-designed distributed systems can ensure the good behaviour of all participants.

Case studies



Weather Ledger

THE TRUST/DATA EXCHANGE CHALLENGE

A construction site is an incredibly complex web of risk management, contractual agreements, supply chains, plans, errors and more. On top of that, it's all happening outside where bad weather can cause the shutdown of an entire site for weeks at a time. Embedded within contracts are multiple clauses, outlining who is responsible for risk at various times, and compensation events where parties agree there should be additional time or funding made available for completion. There can be significant amounts of money involved in dealing with these clauses and current methods are slow, combative and costly.

HOW WE ARE ADDRESSING THE CHALLENGE

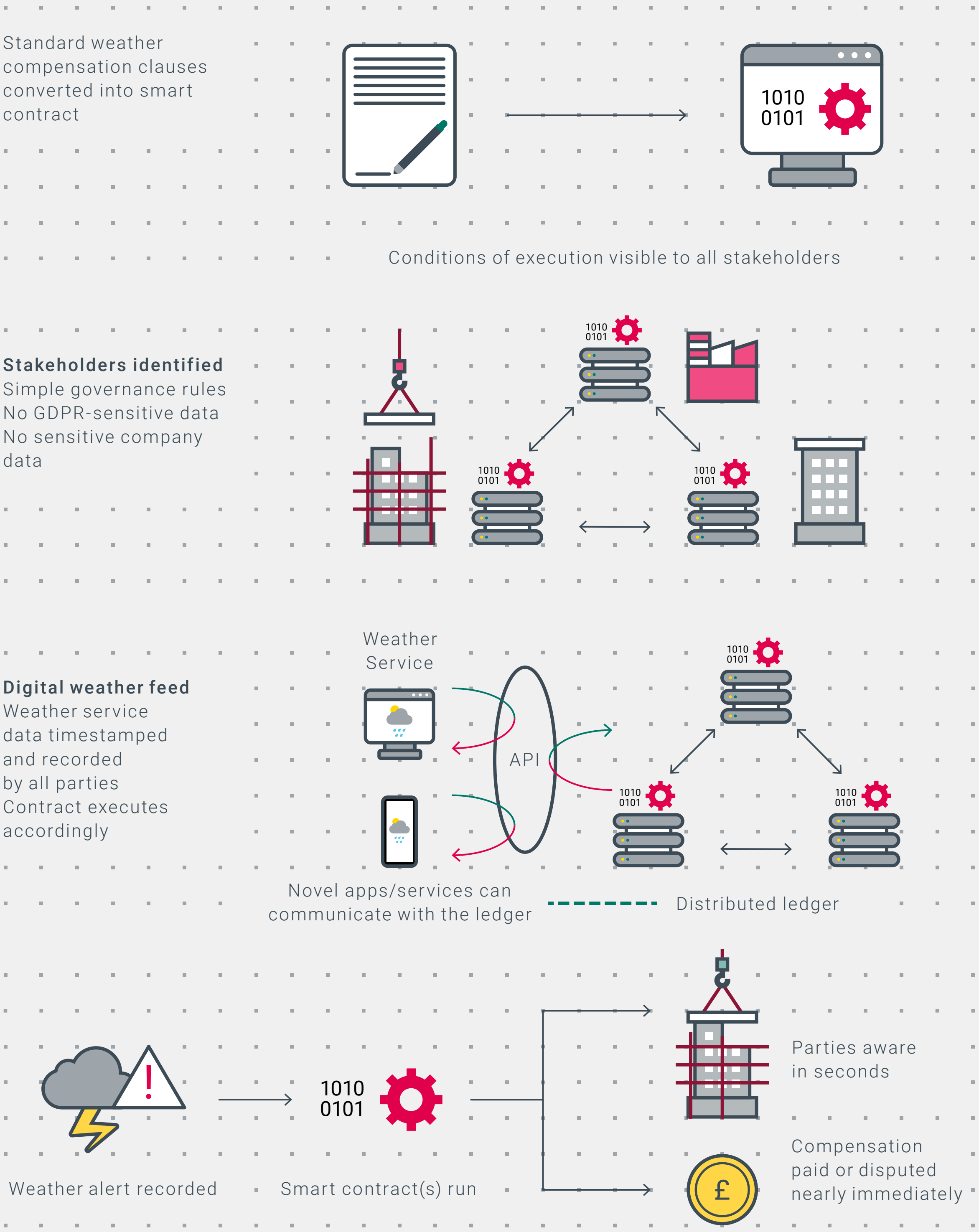
Digital Catapult has formed a partnership with Bam Nuttall, Ferrovial, Connected Places Catapult, Clyde & Co., and Ehabitation Ltd. The partnership recently completed a 12-month project to deploy Internet of Things (IoT) sensors, distributed ledger technology and smart legal contracts to change the way sites manage weather-related risks. This prototype system is called the Weather Ledger, and it has been deployed and tested live on three UK construction sites to date. It proves that this combination of advanced technologies results in a commercially-relevant blockchain solution that will save time, money and tempers in the construction industry.

HOW TRUST/DATA EXCHANGE IS ENABLED

The Weather Ledger represents the first instance of an automatic, legally-compliant smart contract working in an industrial setting. It is a first for blockchain technology and will hopefully lead to further digitisation of contractual clauses that are universally dreaded across the industry. The project explores the reduction in time and cost of weather-related disputes for the construction industry, enabled by a blockchain and IoT-based trusted source of truth for hyperlocal weather conditions. The Weather Ledger feeds the data into smart contracts, compatible with UK law, to create semiautomated, multi-party, legally binding clauses derived from standard NEC construction contracts. The hope in the future is that this approach can eliminate the majority of costly worksite disputes and enable the resolution of compensation events in near real time. This will obviate significant manual claim and claim validation processes and reduce litigation.

The partnership's vision is to make IoT sensors available on every UK construction site. These can feed information into a local ledger that reduces disruption and creates a strong body of data. In turn, this can help with predictive analytics for bad weather events and could be relevant across other sections of the economy. As the climate crisis escalates, we anticipate extreme unpredictable weather to have an even greater impact on construction, which this technological approach goes a long way toward mitigating.

Weather compensation example



Standard weather compensation clauses converted into smart contract

Conditions of execution visible to all stakeholders

Stakeholders identified
Simple governance rules
No GDPR-sensitive data
No sensitive company data

Digital weather feed
Weather service data timestamped and recorded by all parties
Contract executes accordingly

Novel apps/services can communicate with the ledger

Distributed ledger

Weather alert recorded

Smart contract(s) run

Parties aware in seconds

Compensation paid or disputed nearly immediately

Figure 8 - Weather Ledge case study

Pharmaceutical supply chain Field Lab

THE TRUST/DATA EXCHANGE CHALLENGE

The pharmaceutical supply chain for generic medicines is cited as being inherently complex, with lengthy procurement lead times, with multiple stakeholders and a truly globalised nature. For virtual pharmaceutical manufacturers, such as Consilient Health, there is an increasing need to understand the status of assets within this supply chain and also find opportunities to increase its efficiency and transparency of processes.

HOW WE ARE ADDRESSING THE CHALLENGE

We are working with Consilient Health and its global supply chain to determine the feasibility of using DLT systems in four areas:

- **Stock sharing** – The concept of moving from a reactive approach to procuring pharmaceutical batches, to a more proactive and just-in-time approach. Lead times from the raising of a purchase order to the delivery of an approved product to an end-customer can be up to six months (180 days). The majority of this time is taken up in the manufacturing process driven by the contract manufacturing organisation. By sharing batch and forecast data with contract manufacturers, Consilient Health is able to procure raw materials and find space in its factory schedule before a formal purchase order needs to be raised.
- **Logistics tracking** – greater visibility of shipment status would be valuable for all parties involved (contract manufacturing organisation, quality control testing lab and Consilient Health). By combining IoT data (location, temperature, etc.) in a distributed system, batch status information can be visible across organisations to improve timeliness and responsiveness of the supply chain.
- **Digital certificates** – the certification associated with quality analysis and conformity and the burdensome process involved in checking and verifying prior to final release. This is critical to maintaining trust outside of the company and getting batches on the market.
- **Digital identity** – chemicals and pharmaceutical manufacture is often validated by a ‘qualified person’, who is ultimately responsible for signing off batches for release. Currently paper-based, this approval requires a wet ink signature to release the batches to market.

HOW TRUST/DATA EXCHANGE IS ENABLED

By sharing information using a secure, immutable system, Consilient and its supply chain are able to work in a much more agile way. Traditionally, it can take more than six months to manufacture and deliver an order. This is due to closed manufacturing schedules, a lack of visibility of customer requirements and laborious, manual paperwork. If a manufacturer has an improved visibility of the products its customers need, it can create manufacturing schedules to meet these demands.

If a customer is able to track the progress of an order, both in terms of the manufacturing processes and the physical location of the order, new ways of procurement are enabled. For example, this could involve releasing funds when a batch is approved for delivery, leaves the warehouse or enters the final territory.

By digitising certificates and identities, the potential for counterfeit manufacture is reduced and orders are not held up waiting for a specific authority to sign off. Overall, trusted data exchange across the pharmaceutical supply chain improves the efficiency of procurement, manufacture, quality control and distribution systems.

VITALam

THE TRUST/DATA EXCHANGE CHALLENGE

Additive manufacturing (AM) is an emerging technology with high potential for application in the aerospace industry. In metal additive manufacturing, powder quality and the production process have a significant impact on final product quality. Throughout the supply chain, powder is tested and re-tested multiple times, as organisations do not trust the test certificates of the previous organisation. Test certificates are attached to powder batches in paper or PDF formats, but without full backing data. As a result, they are highly susceptible to fraud, which leads to a lack of trust. Improving trust in testing and production data throughout the supply chain could reduce costs and improve delivery times, thereby growing a case for digital certification.

HOW WE ARE ADDRESSING THE CHALLENGE

Digital Catapult is developing a distributed systems infrastructure that will enable materials testing organisations to test and self certify their test data by making it accessible to the entire supply chain. This combines the benefits of a distributed ledger to register critical events within the powder handling lifecycle, and a distributed file system to store the associated large metadata objects for each event.

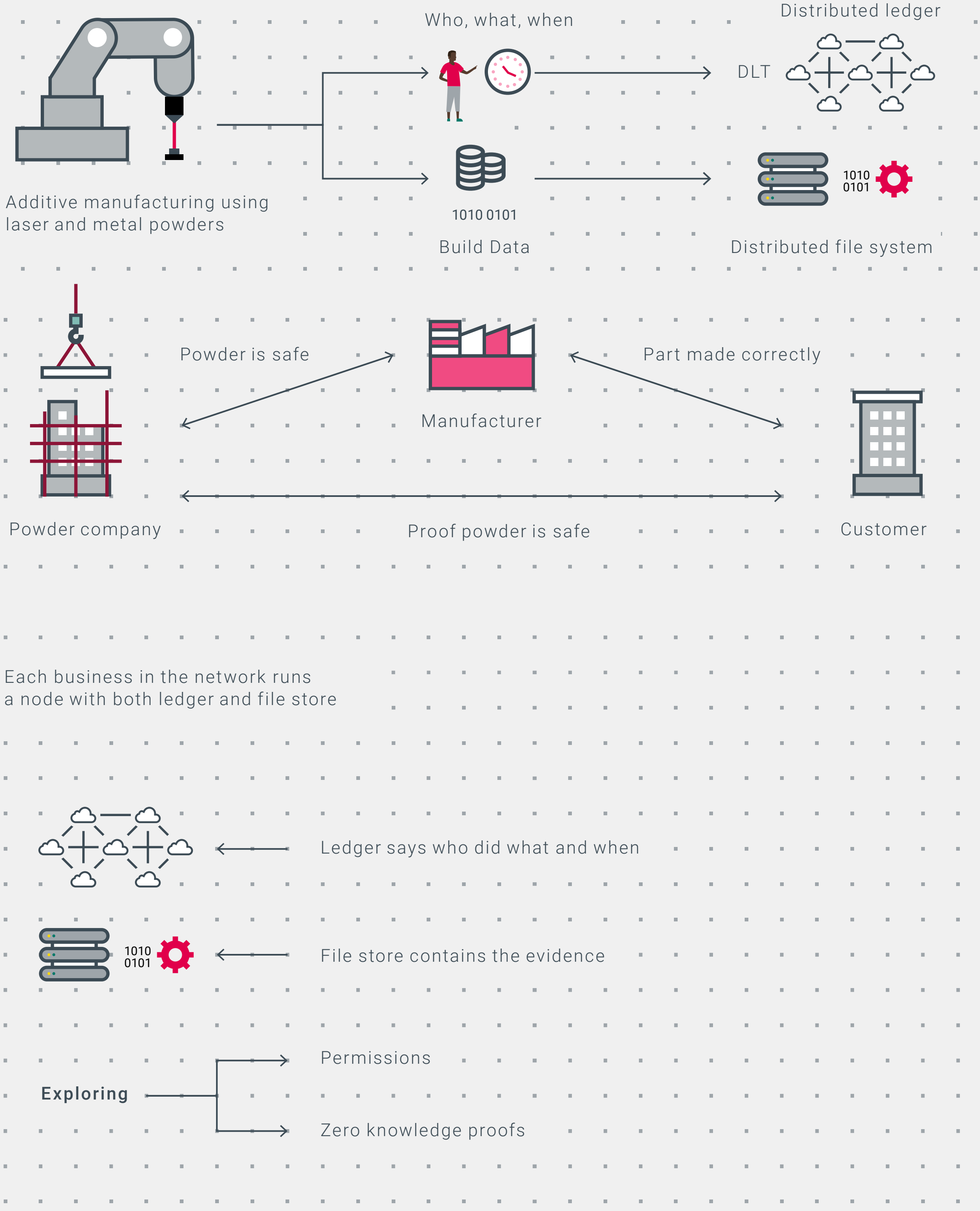
During this project, we interviewed aerospace industry stakeholders throughout the supply chain, including OEM, Tier 1 and small to medium sized enterprise (SME) level personnel to understand the key challenges around trust and data exchange. Powder testing emerged as the most significant challenge. The next activity was to explore the user experience to guide the development of the infrastructure and mock-up graphical user interfaces (GUIs).

HOW TRUST/DATA EXCHANGE IS ENABLED

The project delivered a working prototype demonstrator for the testing documentation use case. This demonstrator allows the organisation that performs the testing to upload the test certificate onto a distributed file system linked to assertions around that event stored in a shared ledger. Test data will become part of an immutable record, accessible to all participants in the supply chain.

> As more test data is shared across the supply chain, it is possible to create new insight into the lifecycle and behaviour of the metal powder.

VITALam example



Who, what, when

Distributed ledger

DLT

Additive manufacturing using laser and metal powders

1010 0101

Build Data

1010 0101

Distributed file system

Powder is safe

Manufacturer

Part made correctly

Powder company

Proof powder is safe

Customer

Each business in the network runs a node with both ledger and file store

Ledger says who did what and when

1010 0101

File store contains the evidence

Exploring

Permissions

Zero knowledge proofs

Figure 9 - VITALam case study

INTELI

THE TRUST/DATA EXCHANGE CHALLENGE

INTELI is a project funded by ISCF Made Smarter, with BAE Systems as the lead partner, with Digital Catapult, Codegate, Maher, AI Labs and Accenture as project partners. The goal of INTELI is to improve the process for supply chain engagement and operations to enable a single source of truth for supply chain data. It aims to optimise scheduling, flexibility and manufacturing capability, as well as make the supply chain more resilient, robust and accessible to SMEs. Fundamental to addressing this challenge is the task of exploring and developing standardised methods of trusted data sharing through a secure, multi-party platform and integrated systems.

The project will focus on hot isostatic pressing (HIP) as the primary manufacturing use case. This is a near-net shape manufacturing process that is difficult to control and significant post-processing is often required to achieve the original design intent. Within the supply chain, there are frequent disputes between the as-designed vs as-built product. By sharing key design and manufacturing process data between buyer and supplier, significant operational and quality improvements can be achieved within the supply chain.

HOW WE ARE ADDRESSING THE CHALLENGE

We are adapting the distributed architecture built for VITALam to the needs of this challenge and the use cases defined in the project. This will involve capturing CAD and design data, as-built data and the differences between these, then tracking the built component through to installation.

The following steps will be taken:

- Build distributed architecture for project challenges/needs for example, re-engineering of VITALam system for the as-designed vs as-built use case
- Define and develop the user experience for distributed systems toolset
- Mock-up GUIs based upon user profiles and market penetration of existing supply chain software solutions
- Design a proof-of-concept decentralised federated learning system operating on data within the ledger

At the end of this project there will be an open source distributed system for forecasting that operates without any coordinating central server, resulting in a trusted method of data exchange.

HOW TRUST/DATA EXCHANGE IS ENABLED

The project focuses on the capability to seamlessly transfer data between organisations and enable visibility and trust within an organisation's manufacturing systems. Its goal is to drive real-time decisions for scheduling, process optimisation and manufacture. The project also aims to innovate solutions to non-technical issues, including suppliers' digital readiness, data trust, data provenance and secure collaboration for data access security and segmentation.

A shared state of knowledge between disparate stakeholders is critical to any collaborative enterprise such as supply chain management. By applying a modern technological approach, this can be achieved without a central controller. Proof of data existence can be separated from data availability and from the visibility of raw data. The technologies to enable these aspects of knowledge and data exchange will be explored in ways that make the most business sense to this supply chain.

Through data trust and sharing, the system's ability to optimise manufacturing and production schedules will be significantly enhanced. This is made possible through visibility of organisations' operations data and in-process monitoring across the supply chain hierarchy. Demand forecasting enables predictive analytics to foresee customer needs and dynamically schedule product supply to the linked supply chain. This enables suppliers to optimise their stock inventory and reduce product lead times, allowing end users to manage stock through advance forecasting/scheduling.

Sellafield: Nuclear ledger

THE TRUST/DATA EXCHANGE CHALLENGE

We are investigating the use of a distributed systems approach creating a secure shared industry ledger to act as the single source of truth for managing low level nuclear waste (LLW). Most LLW comes from the operation and decommissioning of nuclear facilities and can include scrap metal, paper and plastics. Smaller amounts of LLW also come from hospitals and universities. About 94% of all radioactive wastes (by volume) are in the LLW category. It therefore represents the biggest overall challenge for waste tracking and management.

LLW is generated, handled and stored by a number of different organisations in the UK. The sheer volume of waste and its generation over a long time frame creates a huge challenge for hand-off and traceability. Although a wide range of record systems have been employed over the years, these have been abandoned as technology has evolved. As a consequence, a consistent record, if it existed at all, would be dispersed across each of the different systems - starting with paper, through to microfiche, magnetic media and eventually electronic databases - both local and online.

The diversity of data sources, multi-stakeholder nature and evolution of storage media presents a major barrier to the task of understanding what the different LLW stores actually contain and where specific items of waste are located.

HOW WE ARE ADDRESSING THE CHALLENGE

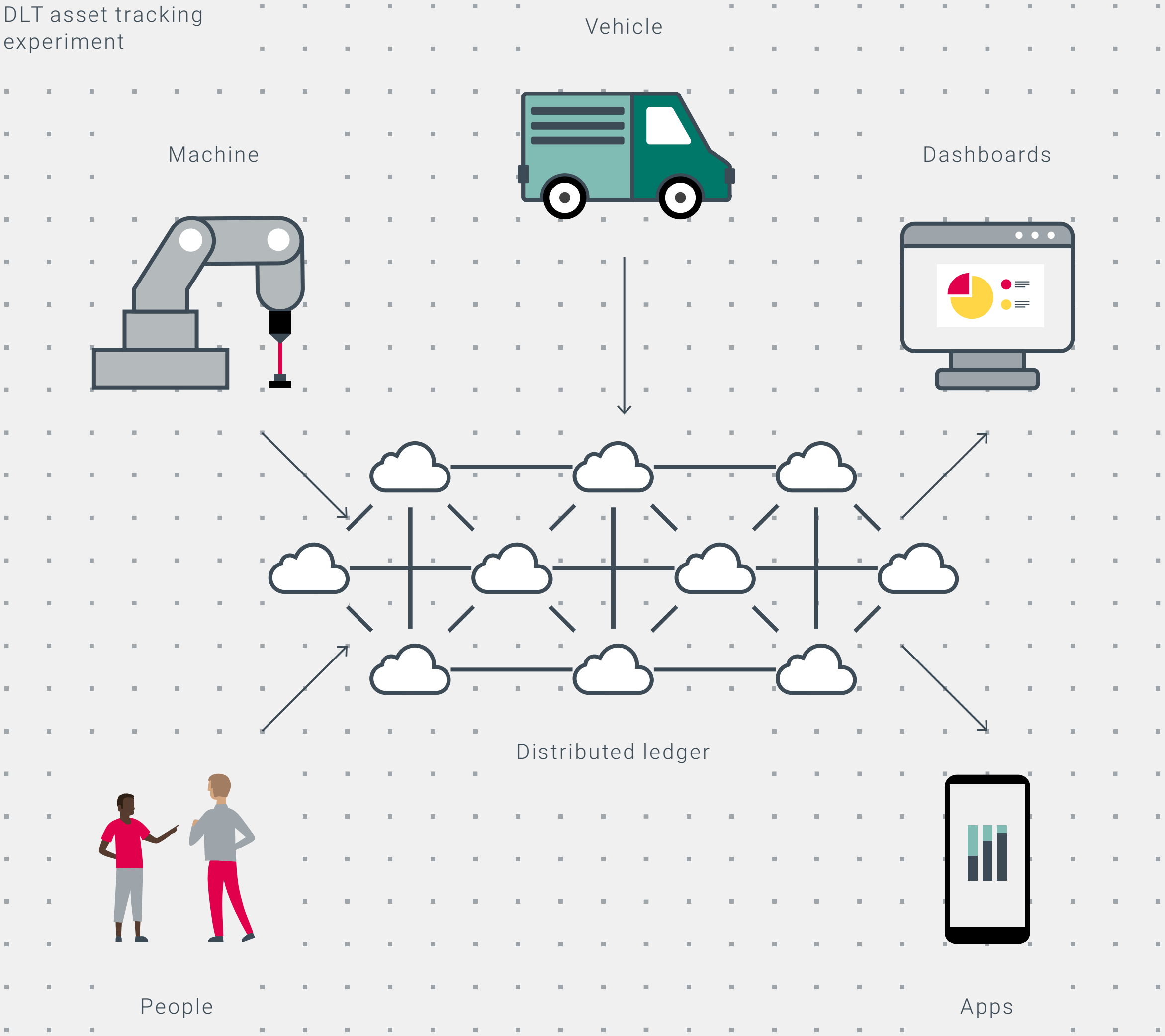
At Digital Catapult, we have developed an in-house approach termed DLT Field Lab. It begins with workshops to identify and prioritise the key challenges in the organisation (or industry) which could be most effectively addressed through the use of distributed systems technologies. Once these challenges are agreed, we help identify potential technology partners through an open call to the DLT innovator ecosystem. A shortlist of the most suitable companies are then invited to pitch their ideas, credentials and skills. The selected company is then contracted to create a working version of a DLT solution to be trialled in near-real-world conditions. Over a period of several months, this will be loaded with live data to demonstrate the effectiveness and value of the solution. Critically, a DLT Field Lab results in not just proof-of-concept but proof-of-value.

HOW TRUST/DATA EXCHANGE IS ENABLED

Once the DLT approach has been demonstrated to solve the challenge effectively, a scaled and fully functional system can then be designed and built to incorporate all the learnings from the Field Lab.

All organisations within the supply chain can now be connected to the shared ledger. Data related to the LLW batches will be added and verified as it is generated. It will be accessed securely, with each partner playing a role in the registration and tracking of new waste. This immutable and secure ledger will now be the highly resilient, multiply-redundant authority for the location and condition of all materials passing through the system. It will be a key enabler in the UK's programme to identify a long-term geological storage solution for nuclear waste.

Sellafield example



All data recorded on shared ledger
- Digital sensors and manual entry
- Verification replaces trust
- Potential process automation
- Immutable, auditable, rich data
- Live, redundant, and accessible

Figure 10 - Sellafield nuclear ledger case study

Sellafield: Nuclear skills passport

THE TRUST/DATA EXCHANGE CHALLENGE

We are planning to improve the productivity of the nuclear sector by facilitating a more mobile and efficient workforce. The challenge is around the verification of skills, qualifications and the site knowledge of workers such as welders, fitters and engineers within the industry.

To start work on a new site, workers are required to be able to demonstrate that they have the necessary training and local knowledge in order to work safely. Verification of existing paper credentials can take time and the accuracy is often questioned. It is generally easier to send new contractors or employees on induction and orientation safety courses before they can start work, whether this is necessary or not. This reduplication of work wastes worker time, site time and money.

HOW WE ARE ADDRESSING THE CHALLENGE

The DLT Field Lab process is identical to the one followed for the Nuclear Ledger.

HOW TRUST/DATA EXCHANGE IS ENABLED

The system created is an early industrial example of a self-sovereign identity solution mentioned above. Verified training records, safety certificates, and on-the-job assessments of competency, etc. are entered onto a person's skills profile in a secure ledger. They are then verified by the agreed consensus mechanism and made permanent and tamper proof.

The transactions are also transparent and fully traceable – allowing analysis of how effectively and accurately they are assessed on different sites.

On arrival at a site, the worker now has control over their identity and can prove all the required and relevant skills and experiences to be instantly verified by the employer. The person is therefore able to commence work immediately without the need to carry any supporting data about themselves.

The system will create a wider record of all the skills available and deployed at the site and ensure that they are adequate to meet industry regulations and safety requirements. It will also identify trends, bottlenecks and predict areas of skills shortages for the longer term benefit of the sector.

Sellafield example

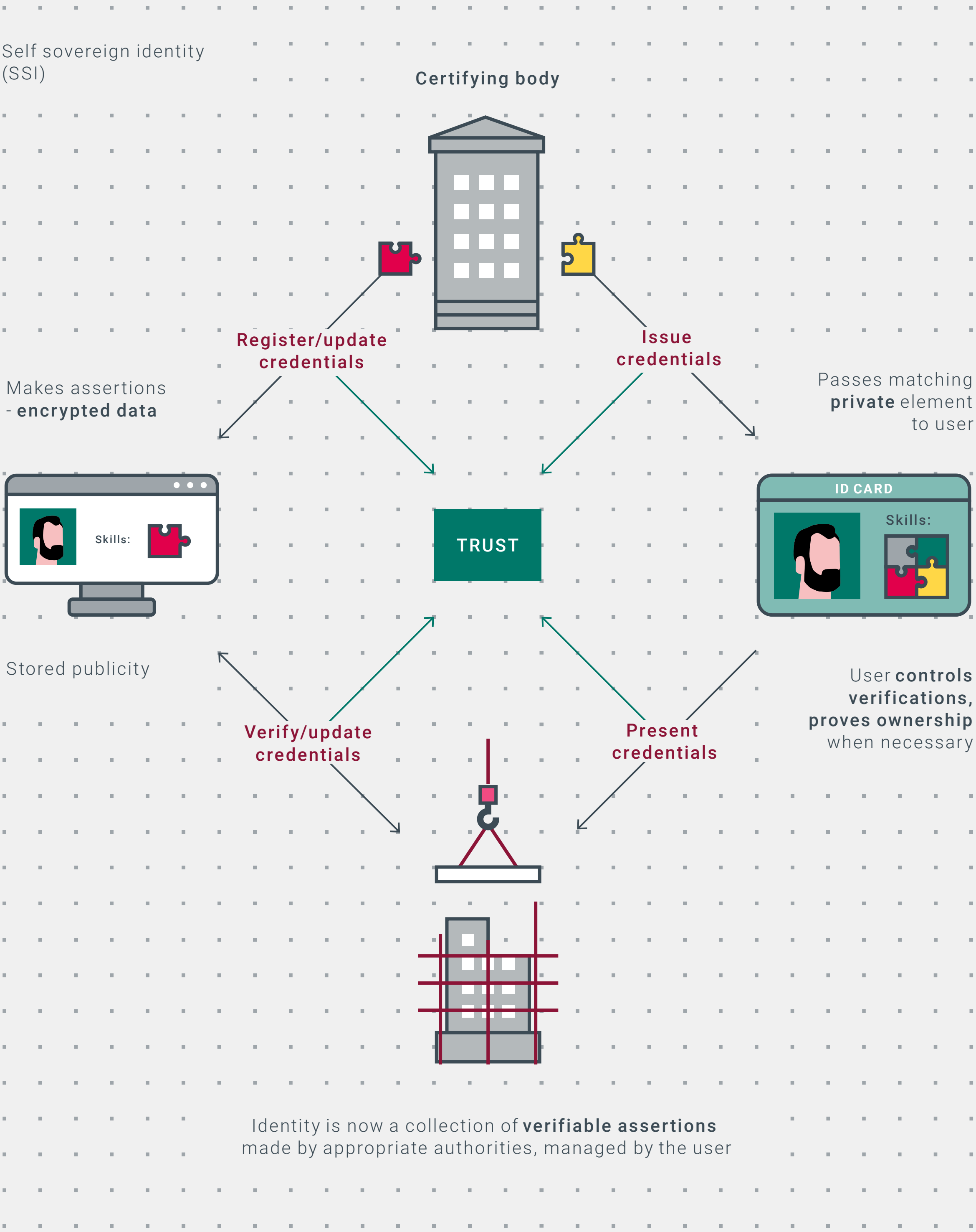

Self sovereign identity (SSI)
Certifying body
Register/update credentials
Issue credentials
Makes assertions - **encrypted data**
Passes matching **private** element to user
Skills:
TRUST
ID CARD
Skills:
Stored publicity
User **controls verifications, proves ownership** when necessary
Verify/update credentials
Present credentials
Identity is now a collection of **verifiable assertions** made by appropriate authorities, managed by the user

Figure 11 - Sellafield nuclear skills passport case study

Digital Sandwich

THE TRUST/DATA EXCHANGE CHALLENGE

Digital Sandwich is a £10 million project, with 40% funded by Innovate UK and 60% from a consortium of ten partners. Led by Raynor Foods, a leading UK sandwich maker, Digital Sandwich is exploiting advanced digital technologies to revolutionise the UK's supply chain for ready-made food.

In June 2019, six hospital patients died after eating listeria-infected sandwiches, but it took over one month to identify the source of the infection and shut down the production line. The sandwich company that made the infected products was then forced into liquidation. This tragedy is a clear example of the consequences of poor data sharing across a sub-optimal supply chain. Digital Sandwich is half way to resolving this problem.

To demonstrate the benefits and value of the digitalisation of the food supply chain, the Digital Sandwich project will create an open multi-party software platform connecting organisations across the ready-made food supply and value chain, from primary production to retail. Ultimately, the demonstrator created by the project will be accessible to thousands of SMEs in the UK's food supply chain for whom this technology would previously be out of reach.

Some of the key priorities for data flow:

- **Collaboration and sharing information** to help reduce the negative impacts of shocks and strengthen supply chain adaptability, flexibility and resiliency in the long run
- **Increased transparency and visibility** to monitor the flow of products, money and information across the entire supply chain, delivering a virtual audit and end-to-end continuous assurance capabilities
- **Supply chain optimisation**, including artificial intelligence (AI) models to optimise supply chain inventory, improve services and reduce waste. Additionally, this will lead to the standardisation of commercial terms, financing and settlement, enabling the complete automation of data across supply chains

Further benefits are realised by unlocking data flow to reduce food waste by 10%, increasing productivity by 10% and reducing the cost of capital and inventory for producers and retailers by at least 2%. Consumers will benefit from the increased safety and traceability of ready-made food, improving public health and public trust in the food supply chain. It will enable immediate identification of provenance and the exact processes that could cause the creation of faulty products.

HOW WE ARE ADDRESSING THE CHALLENGE

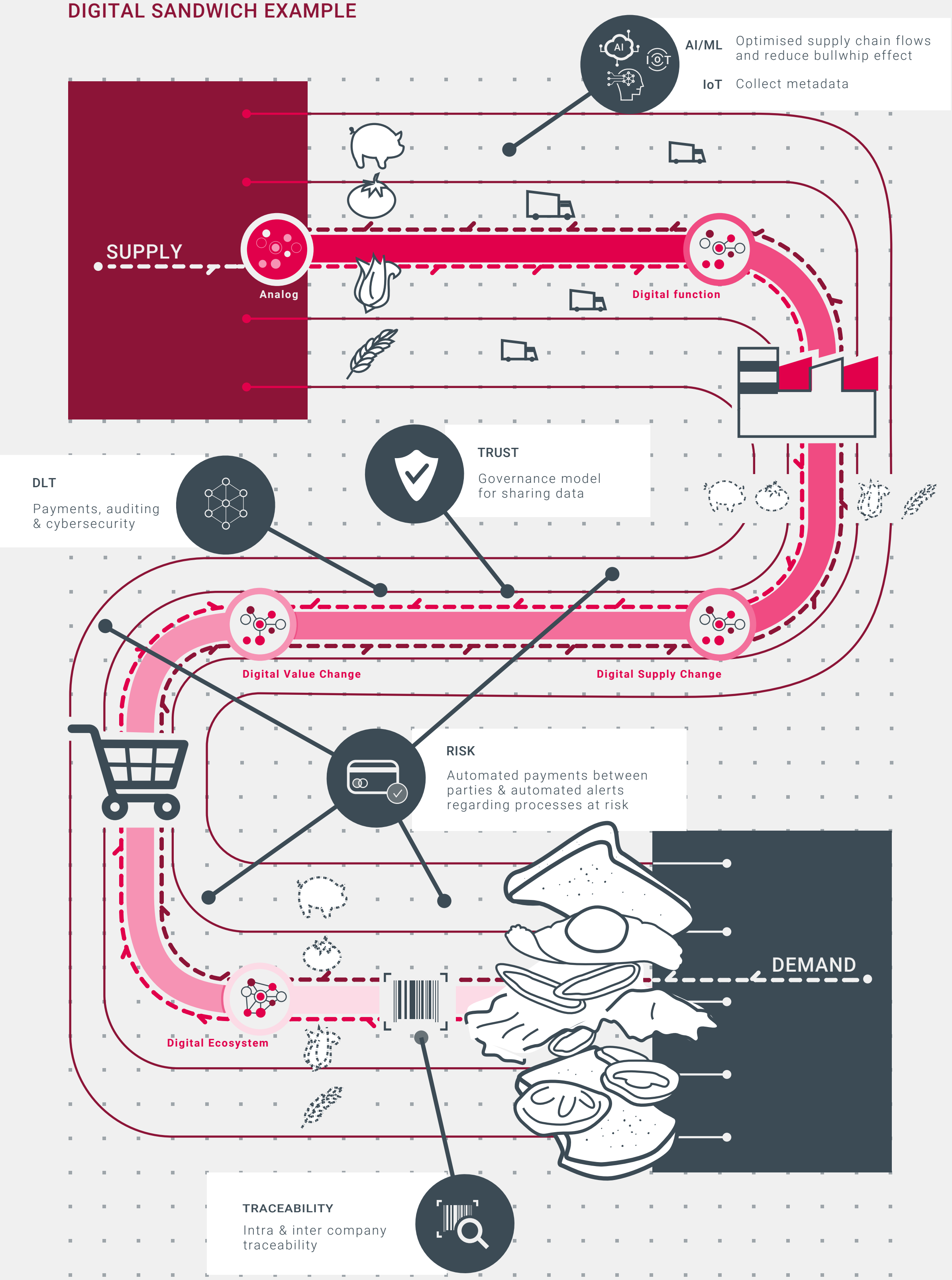
In order to resolve the governance challenge and ensure all stakeholders are comfortable with information sharing, we are using the FSA Data Trust framework. Consortium members have been central in the design of the framework and it is being piloted in the project. This will mean that the value chain community who participate in the project will co-create the rules, gaining confidence that the system has their interests at heart.

To maintain the interests of each stakeholder, we are ensuring that the technological infrastructure supports the appropriate permissioning rules and that the computable contracting functionality will be executed in line with the data trust standards.

HOW TRUST/DATA EXCHANGE IS ENABLED

The governance rules will be complemented by an incentive system developed by one of the consortium partners to maximise the impact of savings made by using the system. This will mean that all stakeholders who actively use and promote the participation of their supply chain partners will benefit financially from their participation. This system will be the engine to incentivise data sharing at scale across participating supply chains. It will drive the network effect that is needed to establish the programme as a national demonstrator of the digital supply chain, first in food and then across other supply chains in the UK and internationally.

DIGITAL SANDWICH EXAMPLE



AI/ML Optimised supply chain flows and reduce bullwhip effect

IoT Collect metadata

TRUST

Governance model for sharing data

DLT

Payments, auditing & cybersecurity

RISK

Automated payments between parties & automated alerts regarding processes at risk

TRACEABILITY

Intra & inter company traceability

Figure 12- Digital Sandwich case study

Glossary

A

Airdrop

A distribution of tokens free of charge to the cryptocurrency wallets of certain users, with or without advance notice. Typically carried out to reward loyal users, or create a buzz about a particular token.

Altcoin

A term used to describe cryptocurrency alternatives to Bitcoin such as Litecoin and Ether. Altcoins can arise from forks of the Bitcoin software code, but many innovative Altcoins (for example Ethereum, Monero, NEO, Polkadot, and others) use completely separate computer code and blockchains from Bitcoin.

Asset tokens / asset backed tokens

These tokens represent assets such as a debt or equity claim on the issuer – for example, a share in future company earnings or future capital flows – and may be tradable as investments. Asset backed tokens reflect an underlying physical asset such as gold.

ASIC

An Application Specific Integrated Circuit – a silicon chip dedicated to only one task, such as performing the hash algorithm used to secure a proof-of-work blockchain (Bitcoin is now only mined with ASICs).

B

Bitcoin (BTC)

The best known cryptocurrency created by Satoshi Nakamoto is a cryptocurrency facilitated by a blockchain.

Bitcoin maximalist

A person or entity that believes only the Bitcoin cryptocurrency deserves to survive long-term out of all cryptocurrencies on the market.

Block

Packages of data recorded on the blockchain, providing important information about the block, including its hash, the hash of the previous block, the nonce, timestamp, the difficulty and the block reward.

Block height

The number of a given block, counted from the genesis block.

Block reward

The reward a miner receives for each new block that it mines (in the case of the Bitcoin blockchain, 6.25BTC).

Blockchain

A peer-to-peer, decentralised, immutable and distributed ledger which consists of validated blocks linked into a time-sequenced chain (Imagine a spreadsheet which is not operated by a central party but is operated by a network of computers and is designed to constantly refresh and update its content so that new entries are time stamped and seen by all operators). The best known blockchain is Bitcoin.

C

Chain tip

The most recent block added to a growing blockchain.

Cold storage

The storage of cryptocurrency offline to protect it from hacking, e.g. in a USB drive or hardware wallet.

Consensus

Computational agreement between nodes in a DLT that the current state of the shared ledger is mathematically valid.

Corda

An open source permissioned DLT created by R3.

Cryptoasset

A digital asset. Includes cryptocurrencies and tokens. Depending on the particular characteristics of the cryptoasset it may or may not be a regulated product.

Cryptocurrency

A form of digital money that is exchanged via DLT. The most widely known is Bitcoin. (A number of policy makers do not consider Bitcoin or other cryptocurrencies to meet the requirements of money in the traditional use of the word and so prefer the term cryptoasset.)

D

DAO (decentralised autonomous organisation)

A fully automated business entity that operates without human involvement. The 'organisation' acts in accordance with its rules which have been transformed into software code. Smart contracts are programmed to carry out certain activities on behalf of the DAO. The best known DAO is the DAO which was built on Ethereum and intended as a new form of venture capital fund. It failed after a vulnerability in the DAO's code led to the equivalent of $70million being siphoned from the fund.

dApp (decentralised application)

A software application created to run on a DLT-based system. Ethereum, EOS, and Polkadot are popular open permissionless DLT platforms for creating dApps.

Difficulty

Mining a block is difficult because the hash of a block's header must be lower than or equal to the target hash in order for the block to be accepted by the network. Put simply, the hash of a block must start with a certain number of zeros:

000000000000000000004e2123bdbd354a87cd51e176a2d3235e3d30ebd20045 – this is the hash of a Bitcoin block mined on 13 July 2018 - as you can see it has 20 zeros. Therefore, the probability of randomly selecting a nonce value which results in a hash that is less than or equal to the target value is very low, and therefore many hundreds of billions of different nonce values need to be tested.

D

Digital signature

A digital code that is created and validated by public key encryption, proving that only the holder of the private key could have generated the signature. This can be attached to a document sent electronically to identify the sender of the document, without revealing the sender's private key.

Distributed ledger technology (DLT)

A distributed and cryptographically secured 'database'. Blockchain is one type of DLT.

Double spending problem

Given that digital information is so easily reproduced, once a digital currency is spent, how do you record this definitively? How can you prevent it getting spent more than once? This had been a longstanding concern with digital currencies. Blockchain technology was invented to prevent double spending without requiring a central trusted authority such as a bank.

E

ERC20 standard

A specific set of functions which developers must use in their Ethereum tokens to make them compliant with a widely established set of protocols and tools.

Ether (ETH)

Is the native token used to operate the Ethereum platform. Ether provides the incentive for nodes to validate blocks on the Ethereum network which contain the smart contract code.

Ethereum

An open-source, public, blockchain-based distributed computing platform released on July 30th 2015 by Vitalik Buterin featuring smart contract functionality that allows developers to build and deploy decentralised applications (dApps).

F

Fiat currency

Money declared by a government to be legal tender (e.g. GBP or USD).

Fork

A Fork is the creation of an ongoing alternative version of the blockchain, by creating two blocks simultaneously at a given block height. Forks occur naturally when two blocks are found simultaneously by competing miners. These types of forks resolve automatically when the next miners choose to build on top of only one of the branches formed. Forks may also be used to intentionally create a new set of rules governing the validity of blocks in a blockchain. See hard fork and soft fork.

G

Gas

Gas is a measurement roughly equivalent to computational steps for Ethereum. Every transaction on Ethereum is required to include a gas limit and a fee that it is willing to pay per gas. Ether miners have the choice of including the transaction and collecting the fee or not.

Gas limit

Is the maximum amount of units of gas the user is willing to spend on a transaction. The transaction must have enough gas to cover the computational resources needed to execute the code. All unused gas is refunded at the end of the transaction.

Gas price

Is the price a user is willing to pay for a transaction in terms of GWei.

Genesis block

The first block of a blockchain, it is generally hardcoded into the software of the applications that utilise its blockchain.

GWei

Each Ether is divisible into 1018 sub-units, called Wei. 1 GWei = 1 gigaWei = 1 billion Wei, or 1 billionth of an Ether.

H

Hard fork

A fork that can render previously invalid types of transactions valid and vice versa. This type of fork requires all nodes and users to upgrade to the latest version of the protocol software. Therefore, a hard fork is a permanent change to the rules of the previous version of the blockchain, and nodes using the previous version will not recognise the new version. A hard fork may be implemented to correct security vulnerabilities, add new functionality, or reverse transactions (see The DAO). Bitcoin Cash is a hard fork of Bitcoin.

Hash

An identifier for input data which does not disclose information about the data. In essence a hash function takes input data and returns a fixed length value which acts as a 'digital fingerprint' for the input data. The hash will always be the same for the same input data. Modifying the input data even by a tiny amount will change the hash in an unpredictable manner. The consensus process securing the Bitcoin blockchain relies on data being hashed using the SHA-256 hashing algorithm.

Hyperledger

An open-source collaborative effort created by the Linux Foundation to advance cross-industry blockchain technologies. This houses multiple distributed ledger projects, including Fabric, Burrow, Iroha and Sawtooth Lake.

I

Initial coin offering (ICO)

An innovative form of crowdfunding. In an ICO, or token sale, a company sells digital tokens that are issued through DLT, typically in exchange for Ether or other cryptocurrencies. In a token sale, the tokens can perform different functions, for example, tokens may take the form of payment tokens, utility tokens or asset tokens.

K

Keys

Public key cryptography uses public and private keys to encrypt and decrypt data. In the context of cryptocurrencies and more specifically Bitcoin, a private key is a secret number that relates to a user's bitcoin address. The private key enables a user to spend Bitcoins as it generates a digital signature, mathematically confirming the user has the right to issue each transaction they send out. The Bitcoins are sent to another user's public key address and become their property, because their private key cannot be calculated from their public key.

L

Ledger

A database which records transactions. In the context of Bitcoin each cryptocurrency transaction is recorded on a public blockchain that is accessible by anyone.

M

Mining

The action of securing a blockchain system using a mathematical process, such as proof-of-work, proof-of-stake or other such methods. Computers who solve these mathematical problems are known as 'miners'.

Monero

Like Ethereum, a completely unique blockchain-based cryptocurrency which is not a fork of Bitcoin. The primary goals of Monero are financial privacy and complete token fungibility, which it achieves by using cryptographic techniques to prevent traceability analysis on its blockchain.

N

Nick szabo A computer scientist credited with coining the term smart contract.

Node Any computer that connects to the DLT network. Nodes first connect to the network and obtain an up-to-date copy of the ledger. Each node is responsible for receiving, validating and relaying transactions and blocks to its peers. This security model (massive redundant distribution with mathematical validation by each participant) ensures permanent availability of data across the network and rejection of invalid transactions.

Nonce In cryptography, an arbitrary number, used once. The nonce is an important concept in proof-of-work mining, as used by Bitcoin, for example.

O

Off-chain Activity that happens, or data that is stored, outside the blockchain ledger, but may be referenced from it.

Oracle A trusted off-chain agent for a distributed ledger system which can submit information to be used by on-chain smart contracts. For example, an Oracle might link to a third party verified source of weather data, travel timetables, stock market information, registry information or to a physical IoT device.

P

Payment tokens Digital tokens which enable the token holder to acquire goods or services from the token issuer (i.e. performs as virtual currency).

Permissioned A DLT system where only pre-authorised nodes can finalise transactions into the ledger.

Permissionless A DLT system where all nodes can access, submit and be selected to finalise transactions into the ledger.

Polkadot A public permissioned ledger built from the ground up and designed to coordinate and enable communication and data relay across multiple independent DLT-based ecosystems. DOTs are the native token of the ledger. Staking these tokens grants permission to take part in consensus.

Private blockchain/ DLT A blockchain/DLT that is only accessible to certain participants. Only pre-authorised nodes can access and submit transactions or finalise transactions. A private blockchain/DLT is always a permissioned blockchain/DLT.

P

Private key

A unique number that acts as a personal password to access cryptoassets in a specific wallet. The key is kept hidden from anyone but the owner of the wallet. Whoever has access to the private key effectively owns the cryptoassets.

Public blockchain/DLT

A blockchain/DLT system which permits anyone with a computer to create a node. A public blockchain can be permissioned or permissionless.

Proof-of-stake

An alternative to proof-of-work. Mining requires a lot of computing power which translates to high electricity usage. Proof-of-stake seeks to address this by limiting what you can mine to the stake of the particular cryptocurrency that you own. (For example, if you own 1% of all Ether available, then you can only mine 1% of the blocks. This also mitigates the risk that miners create competing forks because this would devalue each miner's stake.)

Proof-of-work

Proof-of-work involves using computer processing power to perform repeated hash operations with different nonce values to find a resulting hash below the required difficulty. Finding such a hash allows the miner to add a block of transactions to the chain tip of a growing blockchain. Miners are incentivised to use their computing resources to mine by receiving block rewards. Because this is difficult and consumes large amounts of electricity, it is an effective way of securing the blockchain from attempted rewriting of history (e.g. to double spend) or breaking consensus.

Public key

A cryptographic key used to encrypt messages. A user can sign data with their private key and anyone who knows the user's public key can verify that the signature is valid. However, encrypted messages can be deciphered only by using the paired private key, which cannot be calculated from knowing the public key. A Bitcoin wallet address is a hashed version of the user's public key.

Q

Quantum computing

Quantum computing is seen as a possible threat to the security of blockchain systems because quantum computers are expected to be able to hack public key encryption systems that make blockchain (as well as many other parts of cyberspace) secure.

R

R3

An enterprise blockchain software company working with more than 200 members and partners across multiple industries from both the private and public sectors.

Regulatory sandbox

A controlled space set up by regulators such as the UK’s Financial Conduct Authority to allow authorised and unauthorised firms to test innovative products, services, business models and delivery mechanisms in the real market, with real consumers.

Ripple

Ripple is a real-time gross settlement system. It operates on a consensus mechanism but is not blockchain based.

S

SHA-256 hash

A ‘Secure Hash Algorithm 2’ function which produces 256-bit long output values. The cryptographic hash algorithm used in proof-of-work mining to secure Bitcoin and many other blockchain-based cryptocurrencies (notably not Ethereum, Monero or Ripple).

Side chain

A blockchain that is connected to a parent (primary) blockchain and allows a user to use the cryptoassets securely within that blockchain, but also transfer cryptoassets to and from the parent blockchain.

Smart contract

The term smart contract is rather a misnomer. A smart contract is not a contract in the legal sense, although it could be used to automate elements of a legal contract. Smart contracts are programmable transactions - computer code that sits in an application layer on top of the distributed ledger and acts as an execution mechanism. When certain conditions are met then the protocols automatically execute a set of instructions.

Soft fork

Unlike a hard fork, a soft fork is a software upgrade that is backwards-compatible, i.e. existing nodes will recognise the new code and still be able to function on the network, but not take advantage of the new features on offer. Because of this reduced functionality, soft forks incentivise those who have not upgraded to upgrade.

Staking

Placing value at risk (denominated in the native cryptocurrency of the chain) to vouch for the final state of the chain in a Proof-of-Stake system.

T

Token

Tokens are digital assets issued in connection with an application that uses an existing (such as Ethereum) and can take a variety of different forms. See asset tokens, payment tokens and utility tokens.

Timestamp

Each block contains a timestamp of when it was created. This provides an indication of when a transaction was added to the chain.

U

Utility tokens

A token which provides users with digital access to an application, product or service (such as frequent flyer programmes).

V

Verification

Transaction verification is a mathematical process of checking that a transaction submitted to a Node is a permitted unique transfer of unspent value (see double spending) and that the correct private key has been used to sign the transaction. block Verification checks additional parameters involved in the consensus process, for example that the block has been correctly mined and has an appropriate timestamp, amongst other checks.

W

Wallet

A software application that stores the user's collection of private keys and communicates with the corresponding blockchain / DLT.

Whitepaper

In the context of ICOs, an informational document that provides details on the philosophy, objectives and technology of a given project or initiative and is released in advance of the ICO to attract investment.

X

XRP

XRP is the native cryptocurrency of the Ripple platform. Unlike Bitcoin, XRP is pre-mined i.e. it was all introduced at its inception.

Z

Zero knowledge proof

A cryptographic method by which one party can prove (the prover) to another party (the verifier) that they know secret information, without revealing the secret information. (For example, by way of analogy, in the case of identity, being able to prove that you are over 21 without revealing your actual age or date of birth.)

51% attack

The blockchain platform is protected from attack provided that honest nodes collectively control more CPU power than any cooperating group of attacker nodes. A 51% attack is a situation where over half of the nodes on a blockchain network are controlled by a single malicious miner or a group of miners and such bad actors manipulate the blockchain to their own end (for example censoring transactions including allowing double-spending. 51% attacks cannot manipulate transactions because it is not possible to forge the signatures which secure transactions.)

Footnotes

[1] Routing individual packets of digital information from sender to receiver through different switching stations along the way, rather than along fixed point-to-point paths. Routing can re-map around failing stations.

[2] Servers are nodes that have connections to a large number of other nodes.

[3] Blockchain is a type of Distributed Ledger Technology (DLT).

[4] Co-opetition is a portmanteau of ‘co-operation’ and ‘competition’.

About Digital Catapult

Digital Catapult is the UK authority on advanced digital technology. Through collaboration and innovation, we accelerate industry adoption to drive growth and opportunity across the economy.

We bring together an expert and enterprising community of researchers, startups, scaleups and industry leaders to discover new ways to solve the big challenges limiting the UK's future potential. Through our specialist programmes and experimental facilities, we make sure that innovation thrives and the right solutions make it to the real world.

Our goal is to accelerate new possibilities in everything we do and for every business we partner on their journey – breaking down barriers, de-risking innovation, opening up markets and responsibly shaping the products, services and experiences of the future.

Visit **www.digicatapult.org.uk** for more information.

About the Henry Royce Institute for advanced materials

We are the UK national institute for advanced materials research and innovation. Our aim is to support and grow world-recognised excellence in UK materials research, accelerating commercial exploitation and delivering positive economic and societal impact for the UK.

Royce is ensuring that academics and industry in the UK's materials community have access to world-class research capabilities, infrastructure, expertise, and skills development.

From future cities and their energy supplies, to computing, manufacturing and medicine, the research and innovation facilitated by Royce has the potential to significantly impact peoples' lives.

With its hub in Manchester and with capability distributed across nine founding Partners, Royce works collaboratively to create real solutions and make a fundamental difference to the UK economy.

This paper was produced by Digital Catapult for the Henry Royce Institute and funded by the Engineering and Physical Sciences Research Council (EPSRC).

This report forms part of a suite of complementary roadmapping and landscaping reports designed to stimulate and drive new advanced materials research in the UK:

Materials 4.0: Digitally-enabled materials discovery and manufacturing

Materials for Fusion Power

Materials for End-to-End Hydrogen

Degradation in Structural Materials for Net-Zero

www.royce.ac.uk

Engineering and
Physical Sciences
Research Council