

AI and Health Data: Scaling Innovation Without Losing Control

Few sectors illustrate both the promise and the complexity of the data economy as clearly as healthcare. Vast volumes of data are generated across hospitals, research institutions and diagnostics providers. At the same time, advances in artificial intelligence offer the potential to transform diagnosis, treatment and patient outcomes at an unprecedented scale.

Investment in health technology and AI continues to accelerate, and the technical capability to build increasingly sophisticated models is no longer in question. Yet progress remains uneven. Despite the availability of data and the maturity of AI tools, the pace of real-world deployment continues to lag behind expectation.

The constraint is becoming increasingly clear. The limiting factor is not the capability of algorithms, but access to trusted, usable data.

The Healthcare Data Paradox

Healthcare is one of the most data-rich environments in the modern economy. It generates continuous streams of patient records, imaging data, genomic information, treatment outcomes and real-world evidence. In theory, this provides an ideal foundation for training high-quality AI models.

In practice, however, healthcare is also one of the most data-constrained sectors when it comes to utilisation. Data is fragmented across institutions, subject to stringent regulatory controls and often difficult to combine at scale. Access for research and commercial use is frequently slow and complex. As a result, AI models are often trained on limited or incomplete datasets, insights are constrained, and innovation progresses more slowly than the underlying potential would suggest.

**This is not a problem of data scarcity.
It is a problem of structure.**

Why Health Data Cannot Be Freely Shared

The constraints that exist within healthcare data are not arbitrary. They reflect the nature of the data itself.

Health data is highly sensitive, deeply personal and subject to strict legal and ethical obligations. It is governed by data protection regimes, shaped by institutional responsibilities and often subject to national controls. Organisations are rightly cautious when considering how and whether to share it.

Concerns extend beyond compliance. They include the protection of patient confidentiality, exposure to regulatory risk across jurisdictions, loss of control over how data may be used, and the potential erosion of commercial value or intellectual property. Reputational risk also plays a significant role, particularly in environments where public trust is essential.

These concerns are not easily resolved through contractual arrangements alone. Once data leaves its original environment, control is inherently reduced, monitoring becomes more difficult, and the risk of misuse increases. The result is a natural reluctance to share data at scale.

Fragmentation as the Default Outcome

The practical consequence of these constraints is fragmentation. Datasets remain siloed within institutions. Collaboration is limited or slow to develop. Similar work is often duplicated across organisations, while the broader insights that could emerge from combining datasets remain unrealised.

Even where there is a willingness to collaborate, legal complexity and governance concerns tend to restrict the scope of what is possible. Trust becomes the limiting factor. Organisations are frequently left facing a difficult choice between sharing data and accepting risk, or retaining control and limiting opportunity. This is the central tension within health data and AI.

From Data Sharing to Controlled Access

Addressing this tension requires a shift in approach. Rather than focusing on the transfer of data between organisations, attention is increasingly turning toward models of controlled access. The question is no longer simply how data can be shared, but how it can be used without being freely distributed.

This represents a fundamental change in perspective. Instead of moving data across organisational boundaries, the model enables interaction with data within a governed environment where access, usage and outputs are controlled.

Such an approach allows data to remain within a secure structure while still supporting analysis, collaboration and innovation.

A Structured Approach to Health Data Governance

Emerging models are beginning to demonstrate how this approach can be implemented in practice.

Within a structured framework, such as the Data Asset Foundations (DAF) model being developed in the Isle of Man, healthcare data can be formally defined, segmented and governed according to its sensitivity and intended use. Datasets may be categorised as identifiable, pseudonymised or fully anonymised, allowing different levels of access and control to be applied in line with regulatory requirements.

Crucially, data remains governed at source. Access is permissioned, usage is defined in advance, and all activity is monitored within the environment. This reduces reliance on external enforcement mechanisms and limits the risks associated with uncontrolled data transfer.

Researchers, AI developers and commercial partners are able to access and interact with data within these controlled environments. They can run models, perform analyses and generate outputs without extracting raw data or creating uncontrolled copies. This enables meaningful use of data while preserving the integrity and control required by data holders.

All interactions are logged, auditable and aligned with regulatory obligations, supporting transparency and accountability throughout the lifecycle of the data.

Why This Matters for Artificial Intelligence in Healthcare

The performance of AI models is fundamentally dependent on the quality, scale and diversity of the data on which they are trained. Improving outcomes requires access to larger datasets, broader representation and more consistent inputs.

Achieving this in healthcare necessitates collaboration across institutions and, in many cases, across jurisdictions. However, this must be balanced against strict regulatory requirements, ethical considerations and the need to maintain public trust.

A structured, governed approach to data access provides a viable path forward. It enables scale without loss of control, supports access without unrestricted exposure and allows innovation to take place within clearly defined governance boundaries. In a sector where trust is critical, this balance is essential.

Balancing Innovation & Control

Historically, organisations have faced a perceived trade-off between enabling innovation and maintaining control over data. Structured governance models seek to remove this trade-off.

By embedding clear rules around access, usage and accountability, it becomes possible to enable collaboration while preserving ownership, rights and compliance obligations. Organisations can participate in broader data ecosystems without compromising their responsibilities or exposing themselves to undue risk.

This represents a shift from risk avoidance to risk management, supported by institutional frameworks rather than informal arrangements.



The Role of Legal & Governance Infrastructure

It is important to recognise that technology alone cannot resolve these challenges.

Effective data utilisation at scale requires legal structures, governance frameworks and clearly defined rights and responsibilities. The Data Asset Foundations model reflects an attempt to bring these elements together, combining legal recognition, structured governance and controlled access mechanisms within a single framework.

Such approaches remain subject to ongoing development and adoption. However, they reflect a broader recognition that data governance is not a secondary consideration. It is central to unlocking the full potential of AI.

Beyond Healthcare: A Wider Application

While healthcare provides one of the clearest examples of this challenge, the underlying principles extend far beyond it.

In any sector where data is sensitive, collaboration is necessary and trust is essential, the ability to structure and govern data effectively becomes a strategic advantage. As industries become increasingly data-driven, the organisations that can enable controlled, trusted data access will be best positioned to innovate.

Conclusion: Unlocking AI Through Governance

The future of AI in healthcare will not be determined solely by advances in algorithms or computing power. It will depend on the ability to access, combine and use data in ways that are both effective and trustworthy.

Without this capability, innovation will remain constrained by fragmentation and risk. With it, the potential for improved outcomes, efficiency and insight is substantial.

The challenge is not simply to share more data, but to create systems in which data can be used, combined and analysed without losing the control that makes it valuable. Structured frameworks such as Data Asset Foundations represent an important step in that direction, offering a model through which innovation and governance can coexist.

Frequently Asked Questions: AI, Health Data & DAFs

Q1 Why is health data difficult to use for AI?

Health data is highly sensitive and subject to strict legal, ethical and regulatory controls, making large-scale sharing and combination challenging.

Q2 What is meant by controlled access to data?

Controlled access allows users to interact with data within a governed environment without extracting or freely distributing the underlying datasets.

Q3 How do Data Asset Foundations support healthcare data use?

They provide a legal and governance framework that enables datasets to be defined, registered and accessed under controlled and auditable conditions.

Q4 Can AI innovation scale without sharing raw data?

Yes. Structured environments allow models to be trained and analyses to be performed without requiring unrestricted access to raw datasets.

Q5 Is this approach limited to healthcare?

No. The same principles apply to any sector where data is sensitive and trust is critical, including finance, energy and public sector data systems.



DataAssetFoundation
DATAASSETFOUNDATION.CO

Data Asset Foundations
10-12 Victoria Street
Douglas
Isle of Man
IM1 2LH

T: +44 (0)1624 639350
E: info@dataassetfoundation.co
W: www.dataassetfoundation.co

