littlefish DATA & AI GROUP | CONSULTANCY INSIGHTS

Tackling Data Sprawl and Classification in Healthcare

Building secure, usable data foundations for the NHS

DATA & AI

When more data becomes a risk

"Across the NHS, data has become one of the most critical tools in delivering safe, effective care. From electronic patient records and diagnostic imaging to remote monitoring, workforce management systems and operational analytics, healthcare organisations now rely on a vast and growing web of digital information to function day to day.

Yet for many Trusts, Integrated Care Systems (ICSs), and health and social care providers, this explosion of data has created a new challenge: data sprawl.

Data sprawl refers to information that is spread across multiple, disconnected systems (including legacy clinical platforms, cloud services, shared drives, third party suppliers, and departmental tools). Ownership is unclear, sensitivity is inconsistent, and visibility is almost certainly limited.

While each dataset may have been created with good intentions, together they can form a fragmented and risky landscape that is difficult to govern, protect or use effectively.

Crucially, data sprawl in healthcare is not just an IT problem; poorly governed data can slow clinical decision making, increase operational burden, expose organisations to cyber risk and undermine public trust. The very information designed to improve care can, without the right foundations, become a liability.

Throughout this document, we'll explore why data sprawl has become such a pressing issue for the NHS, why data classification is central to addressing it, and how healthcare organisations can regain control without slowing innovation or frontline delivery. **I hope you find it useful.**

Mike Wild
Health Chief Technology Officer

DATA & AI

Understanding data sprawl in an NHS context

Data sprawl in healthcare looks very different to data sprawl in other sectors. The NHS operates within a uniquely complex environment, shaped by long system lifecycles, regulatory requirements, and the need to deliver care around the clock.

Clinical data alone flows from multiple sources: electronic patient records, diagnostic systems, imaging platforms, pathology labs, medical devices and (increasingly) also patient generated data from wearables and remote monitoring tools. Alongside this sits a growing volume of operational, workforce, financial and research data – each with its own systems, owners and governance models.

Over time, many NHS organisations have introduced new platforms without fully retiring the old ones. While cloud adoption has brought welcome flexibility and improved collaboration, it has also multiplied the number of places where data can live. Shared drives, email attachments, collaboration tools and supplier portals all add to an ever-expanding data footprint.

Layered on top of this is the legacy of shared Azure tenancies, which provide access to an M365 OneDrive. In some cases, this sits alongside a dedicated tenancy with a separate OneDrive location, secured with different credentials. Rather than simplifying day-to-day work, this overlap often increases user confusion and frustration.

The result can be a fragmented environment where no single team has a complete picture of what data exists, where it resides, or how sensitive it is. This fragmentation makes it harder to secure information consistently, respond confidently to incidents, or ensure data is being used appropriately across the organisation.

Data sprawl is not a sign of failure; it is a byproduct of digital progress. However, without deliberate action, it becomes increasingly difficult to manage.

"Data sprawl is not a sign of failure; it is a byproduct of digital progress. However, without deliberate action, it becomes increasingly difficult to manage."

Nardos Abraham
Consultancy Services Director

DATA & AI

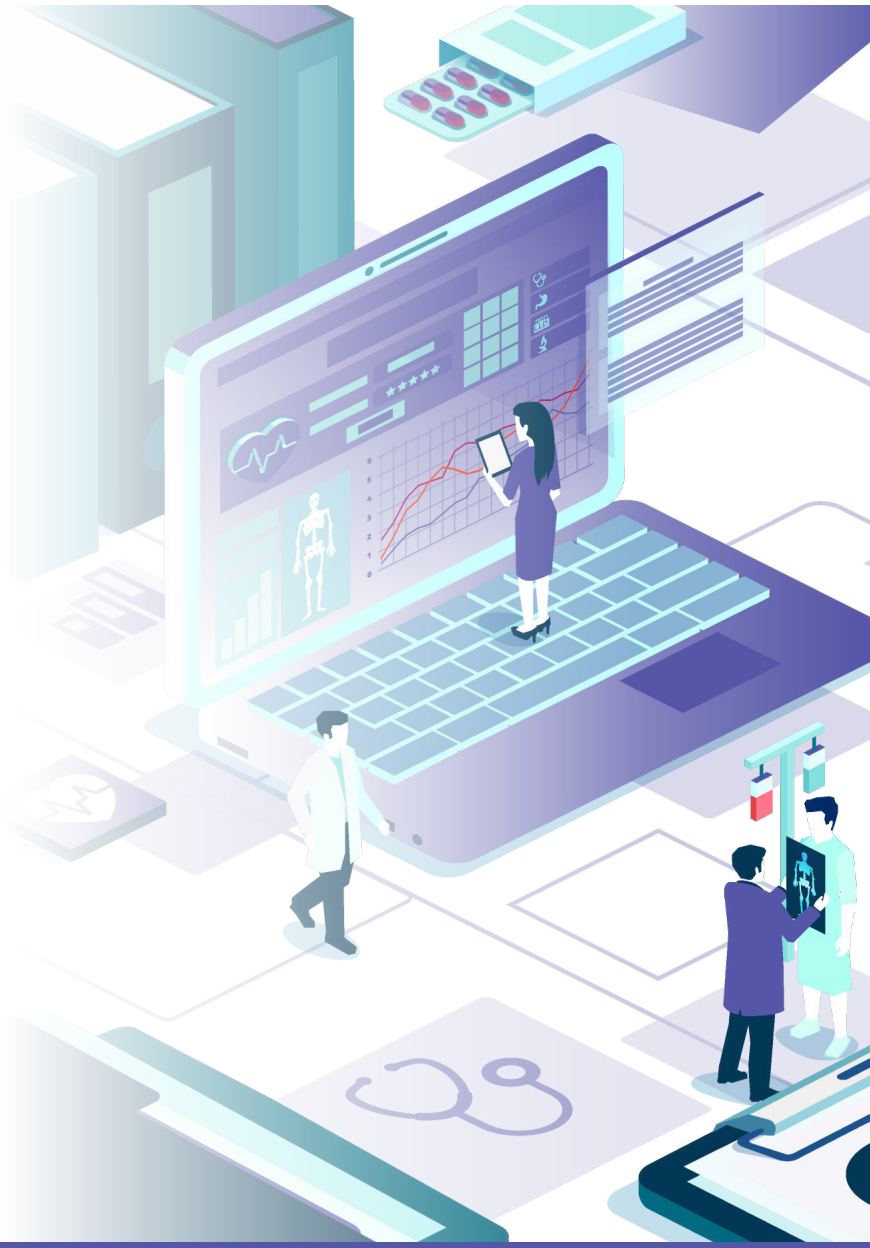
Data classification is the missing foundation

At the heart of tackling data sprawl sits a deceptively simple concept: data classification.

Data classification is the process of understanding what data you hold, how sensitive it is, and how it should be handled. Within the NHS, this means recognising everything from patient identifiable information and clinical risk data to staff records, commercially sensitive material, and operational datasets, and ensuring each is managed in line with its importance and risk.

Many NHS organisations already have well-developed data classification and retention policies in place. The challenge rarely lies with the Information Governance guidance itself, but with applying it retrospectively to data estates that are mixed, fragmented, and historically undefined.

When classification is absent or inconsistent, organisations are often forced into a blunt approach to security and governance. Either everything is locked down (slowing access for clinicians and operational teams) or controls are relaxed, increasing exposure and risk. Effective classification enables a more proportionate and pragmatic model. Highly sensitive patient data can be protected with stronger access controls, monitoring, and retention policies, for example, while less sensitive information can remain accessible and usable. Classification provides the context that allows data to flow safely, rather than blocking it.



"Classification is not about creating more paperwork. Done well, it becomes an enabler for secure collaboration, faster decision making, and confident digital transformation."

DATA & AI

The operational impact of poor data visibility

THE CONSEQUENCES OF UNMANAGED DATA SPRAWL ARE OFTEN FELT MOST ACUTELY ON THE FRONTLINE

Clinicians may struggle to locate the right information at the right time, particularly when data is split across multiple systems that do not integrate cleanly. Administrative teams face duplication, rework, and manual processes to reconcile information held in different places. Analysts spend time cleaning and validating data rather than generating insight.

Over time, this erodes confidence in digital systems. Teams begin to rely on workarounds, shadow IT, or local copies of information to get the job done. While understandable, these behaviours further increase sprawl and risk.

There is also a financial and operational cost. Data that is poorly classified and governed often leads to over retention, inflated storage costs, and difficulty responding to audits or information requests. When organisations lack visibility, they cannot make informed decisions about what data is still required and what can be safely retired.

Where hospitals have implemented a command centre structure, the availability of data reflecting the currently operational activities such as bed state and admissions can allow effort to be focussed where the best operational efficiency will be gained. Historical information, correctly categorised and classified, can help build a predictive picture for the command centre to apply this focus.

Beyond the boundaries of the hospital, identified data with the correct structure applied can be made available via the national Federated Data Platform (FDP) to benefit the local, region and national healthcare systems.

For the NHS, where resources are already stretched, the operational drag of unmanaged data is not sustainable. Clear classification and visibility help ensure data supports care delivery and doesn't slow it down.

DATA & AI

Protecting data without slowing care

Cyber security is now inseparable from the NHS's ability to deliver safe care. Healthcare data is highly valuable, and NHS organisations remain a prime target for ransomware, phishing, and data theft. As digital estates expand, so too does the attack surface.

Data sprawl significantly increases cyber risk. When organisations do not know where sensitive data lives, it becomes harder to protect it consistently. Systems may be exposed unnecessarily. Access may be broader than required. Detection and response become more complex when incidents occur.

Data classification provides a critical foundation for cyber resilience. By understanding which data is most sensitive and business-critical, NHS organisations can apply security controls proportionately and intelligently. This includes aligning access to clinical roles, prioritising monitoring for high-risk datasets, and ensuring backup and recovery strategies focus on what is most critical.

Crucially, this approach supports care delivery rather than obstructing it. Security becomes targeted and context aware, rather than universally restrictive. Clinicians can access the information they need, while the organisation retains confidence that patient data is being protected appropriately.

"In an environment where public trust is paramount, demonstrating control and stewardship of data is as important as preventing breaches themselves."

Sean Tickle
Cyber Services Director

DATA & AI

Governance, compliance, and accountability

The NHS operates under significant regulatory and governance obligations – this includes data protection legislation, clinical safety and information governance frameworks. Yet compliance becomes far more challenging when data is fragmented and poorly classified.

Clear classification supports consistent governance

By establishing ownership, handling rules, and retention policies that can be applied across systems, leaders enable the organisation to respond more confidently to audits, subject access requests, and regulatory scrutiny. All because they understand where data resides and how it is managed.

Importantly, governance is not solely a technical challenge. People and processes play a critical role. Clinical, operational, and digital teams all need clarity on their responsibilities and confidence in how data should be handled. Classification provides a shared language that aligns these groups around common standards.

When governance is embedded into everyday practice, rather than treated as a separate compliance exercise, it supports safer and more efficient use of data across the NHS.

DATA & AI

Act Now: Practical Solutions for Data Sprawl

Addressing data sprawl and classification does not require a wholesale overhaul overnight. In fact, attempting to “fix everything at once” often leads to stalled initiatives and frustration.

Successful NHS organisations take a pragmatic, iterative approach. They start by improving visibility, focusing on high risk and high value data first. They align classification efforts with clinical priorities, cyber risk and operational impact, rather than abstract technical models.

Technology plays an important role, but it must be supported by clear policies and engagement across the organisation. Classification should be designed to fit clinical workflows and operational realities, not imposed as an additional burden.

Over time, incremental improvements compound. Data becomes easier to find, safer to share and simpler to govern. The organisation moves from reacting to issues as they arise, to proactively managing its information landscape.

DATA & AI

Where this leaves you

Data will continue to shape the future of healthcare. From integrated care pathways and population health management to AI-supported diagnostics and virtual care, the NHS's ability to innovate depends on how effectively it manages information today.

Data sprawl and poor classification undermine that ambition. They introduce risk, inefficiency, and uncertainty into systems that must be trusted by clinicians, patients and the public alike.

By investing in clear data classification and governance, NHS organisations can regain control without sacrificing agility. They can protect what matters most, support frontline care, and create a digital foundation that enables innovation safely and confidently.

If you're exploring how to regain control of your data while supporting safe care and innovation, we'd love to continue the conversation.

Get in touch today
0330 390 2002
info@littlefish.co.uk
littlefish.co.uk

"Data doesn't deliver care – people do. And when users trust the systems they rely on, and patients trust how their information is handled, a Trust can move faster and innovate more safely. Getting data foundations right protects the trust and enables better outcomes for all."

Mel Heath
Head of Healthcare