‘Oh! how humanity might embrace the tools that strip away their humanity….’

Martin Blanchard MD BSc MBBS

The ‘new structure’ of the NHS in England is digital (of course). But what has been done to afford it to improve NHS and National productivity? (Part One: the machines and factories; Part Two: the effect on workers)

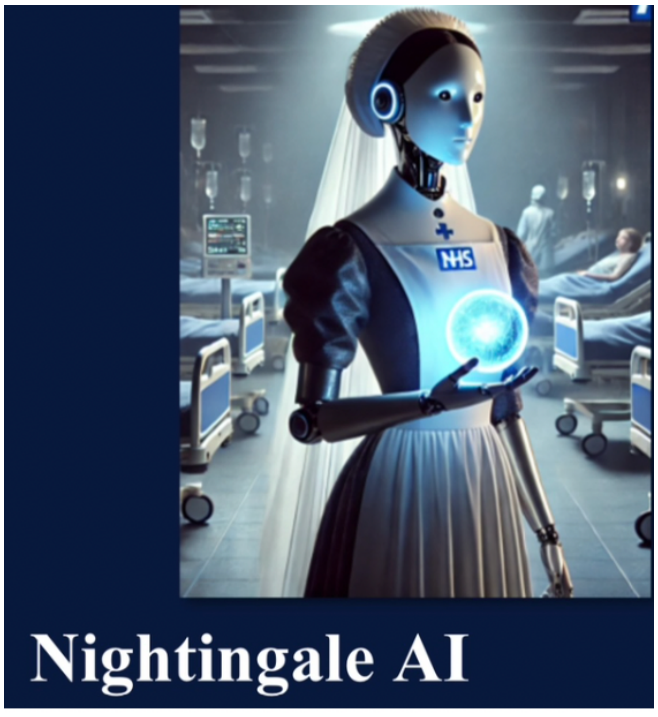
Part one:

The mechanisation of labour within the NHS. How the machine is expected to work.

The arrival of big tech in our NHS is neither fortuitous, nor is it through any sense of beneficence on big tech’s part. It seems that it is instead the latest element of Capital’s and US foreign policy -an enormous expansionary strategy to gather wealth, and to control. All that can be channelled back, or otherwise, is to be used to maintain US hegemony. It is the latest form of ‘imperialism’. We have entered this ‘partnership’ in the hope that we can provide our state funded healthcare system for less money (they call it ‘productivity’) whilst ‘riding on the coat-tails’ of these behemoths so as to gather some ‘super power’ status……The greatest concern is that it is the NHS staff and the ‘least productive’ patients who may have to suffer for all the wealth that is generated.’

The ‘machines’ that underlie the new NHS[1]

(The London Secure Data Environment (SDE) as an exemplar)



[1] In the author’s opinion the most important thing to note is how everything is ultimately dependent upon the use of our data being kept on Microsoft, AWS and Google infrastructure. NVIDIA DGX and Apple and Google (consumer software) have major secondary roles. I WILL TRY TO INDICATE IN THE TEXT WHERE EACH BIG TECH CORPORATION IS INVOLVED AS WE MOVE THROUGH THE WORKINGS OF THE MACHINE/ECOSYSTEM. Immediately, it seems that any major growth in skilled jobs in our economy will be among software engineers! https://d3london.github.io/sde-aic/

Overview

The NHS in England is to depend upon eleven regional Secure Data Environments (SDEs)[2] with a twelfth-the National/Central SDE run by the DoHSC.

- The 11 regional SDEs will be 'infrastructures' for i) the active digitisation and running of digital solutions for their area of the NHS; ii) research and development and iii) the 'commodification' of our health data and the production of 'innovations'.
- The primary aims are, for the NHS as a service to demonstrate increased 'productivity' but also to demonstrate economic benefits nationally. An inseparable aim tied in, is for the private technology providers and their 'partners' to make as much profit from the provision of their services as possible.
- Government has chosen Microsoft, Google, Amazon (AWS), Oracle and NVIDIA big tech corporations as the infrastructure providers for the SDEs rather than to develop any state funded and owned system.[3]
- Big tech have global monopolies in essential cutting-edge compute and storage, AI training data and knowledge, and markets in that they control global digital value chains of production in many sectors of society including increasingly in healthcare.
- Big tech's infrastructures are known as 'cloud' or 'public cloud' and bring with them all the harms to humanity in which they have been involved (see footnotes [4,5,6,7].)
- The SDEs are designed to be **'continuous learning systems'**. They use tech solutions, most of which are based on big tech clouds that can be purchased/rented via hundreds of big tech subordinate 'customer fulfilment' 'partners'/companies. These 'cloud **markets**' are '**asymmetric'** in that big tech control them behind a vague appearance of 'competition'. The SDEs also use i) international, national and local research (mainly US, UK and EU invested), ii) NHS data and iii) the big tech cutting edge 'cloud infrastructure' to generate new knowledge and innovations which, where agreed, can be used to try to improve continuously the NHS performance, and to trade.
- 'Continuous learning' refers not only to humans but also 'AI algorithms' as 'workers'; the latter will become totally embedded in services. All is done knowing that AI, unlike any machines ever created before, will not wear out but instead improve in their performance and value as they use huge quantities of our NHS data.
- Data is the lifeblood of big tech's current business model. It acts as the primary raw material used by them to generate revenue, train artificial intelligence, and

---

[2] East Midlands, East of England, Kent, Medway and Sussex, London, North East and North Cumbria, North West, South West, Thames Valley and Surrey, Wessex, West Midlands, Yorkshire and Humber

[3] Similar choices seem to have been made across our other essential public services

[4] https://www.tni.org/files/202304/Digital%20capitalism%20is%20a%20mine%20not%20a%20cloud.pdf Big tech clouds as mines

[5] . https://www.tni.org/files/2024-09/Digital-Capitalism-Week2.pdf Digital Overlords

[6] . https://www.tni.org/en/article/there-are-no-markets-anymore

[7] Rong, K., Parente, R., Deng, Z. *et al.* The data-based power of big-tech multinational enterprises. *J Int Bus Stud* **57**, 220–235 (2026). https://doi.org/10.1057/s41267-025-00828-5

maintain competitive monopolies. Without the availability of vast datasets, their 'foundation models' would simply collapse[8].

- As well as revenue streams from the infrastructure and software services provided for the NHS, it is the commodification of NHS data for the creation and potential marketisation of 'new knowledge' and innovations (stand alone or systemic) that may be the 'money spinners' on the global market for big tech. This is the '**AI' factory** element of the London SDE which is built around the new 'Foundation World AI Model for Healthcare' within the Artificial Intelligence Centre (AIC) with support from all other areas of the ecosystem and marketed as '**Nightingale AI'**. While the NHS may benefit from the use of innovations, the issue of 'taking' them to market is problematic given our dependence on the use of big tech infrastructure given their monopoly of value chains and markets, and their history of wealth extraction. The huge amount of finance they and their investors have supplied for the future digital NHS means that they will only accept big returns.

The London Secure Data Environment (SDE):

Figure 1: London SDE Map

[8]Foundation models are artificial intelligence (AI) models trained on vast, immense datasets and can fulfil a broad range of general tasks. They serve as the base or building blocks for crafting more specialized applications. A foundation model is itself incomplete but serves as the common basis from which many task-specific models are built via adaptation. The term 'foundation' is used to connote the significance of architectural stability, safety and security: poorly constructed foundations are a recipe for disaster and well-executed foundations are a reliable bedrock for future applications

The London Artificial Intelligence Centre (AIC)

The Artificial Intelligence Centre (AIC) (aka the London Medical Imaging & Artificial Intelligence Centre *for Value Based Healthcare*) has its critical 'cloud' infrastructure provided by Microsoft and AWS and alleged 'sovereign' cloud (see footnote19) provided by NVIDIA and Nebius, an 'out of Russia' Amsterdam based company. In September 2025 Nebius signed a contract with Microsoft worth up to $19.4 billion, and one with Meta worth up to $27 billion in March 2026. That month NVIDIA also took a $2 billion equity stake as part of a 'strategic' partnership.

Nightingale AI is a large-scale, 'foundation world model for health' designed to act like a GPT-style AI, but rather than processing generic text, it digests vast repositories of clinical data, electronic patient records, biomedical data, and life sciences literature.

Data sources: While the exact number of datasets is continuously expanding, the model is being trained on massive 'pools' containing**:**

- millions of electronic health records from the NHS such as routine general practice (GP) records from north-west London, Sussex, and the national Clinical Practice Research Datalink (CPRD)[9]
- data from the US via a major data partnership with the Children's Hospital of Orange County (CHOC)
- unspecified data sources in the EU
- the UK biobank
- biomedical literature and life science databases.

Instead of centralising all this sensitive information in one place, Imperial College has used federated learning to train the algorithm securely across partner hospitals and healthcare systems. The model's development was accelerated by use of the UK's Isambard-AI supercomputer in Bristol using containerised federated Trusted Research Environments (TREs)[10].

Of interest….well at least to some! See foot-note on Synthetic data [11].

---

[9] The database holds de-identified (fully anonymised for researchers) health records for over 60 million patients, including about 16-18 million currently registered patients across the UK. detailed primary care information such as diagnoses, symptoms, prescriptions, immunizations, and referrals. About 25% of the tracked patients have at least 20 years of continuous medical follow-up. CPRD primary care data can be securely linked to 17 other health datasets, including secondary care (Hospital Episodes Statistics), death registrations from the ONS, and cancer registry data. It has contributed to over 2,300 peer-reviewed publications and is used by academic researchers, regulators for pharmacovigilance, and the life sciences industry interested in how treatments perform in diverse, real-world populations.

[10] **Containerised Trusted Research Environments (TREs)** are highly secure digital spaces that use containerisation—like Docker containers—to package research data and analytical tools. They allow approved researchers to analyze sensitive or confidential information (like NHS health records) without the raw data ever leaving the secure server.

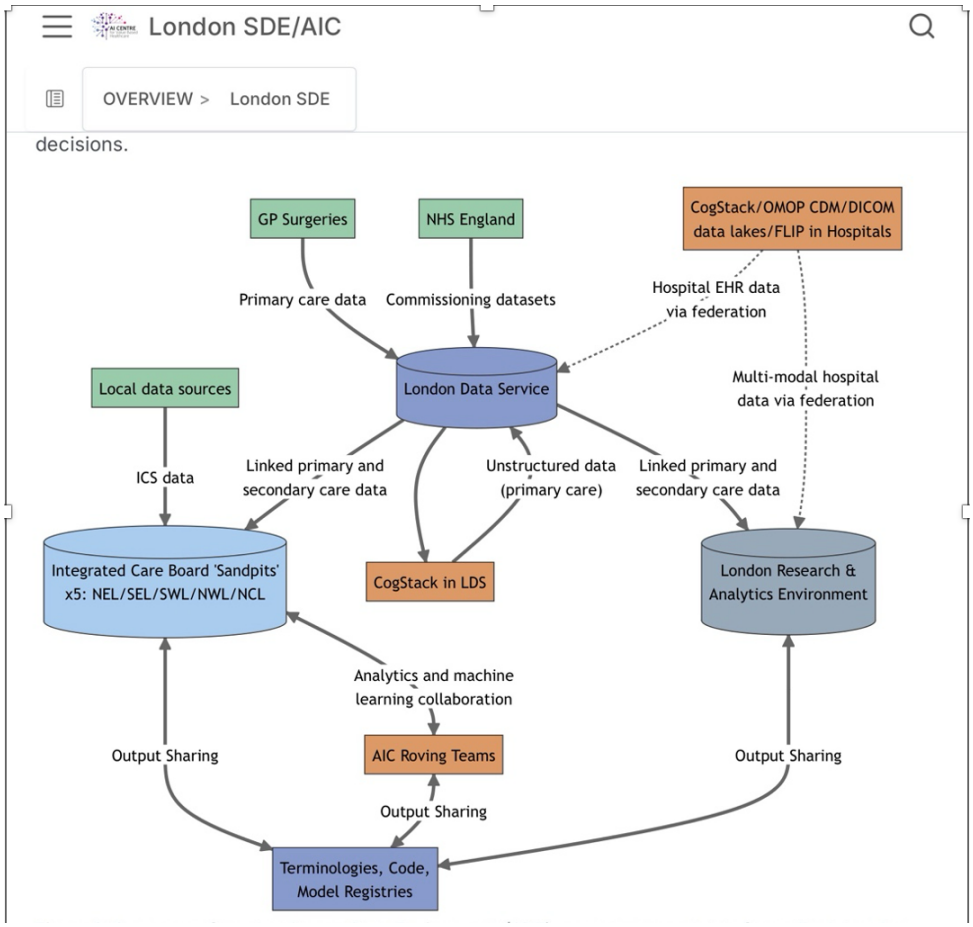
[11] AI models can create synthetic (artificial) data by learning the underlying patterns, statistical properties, and relationships of real-world data. Rather than just copying existing data, models map its structure and generate entirely new data points that are statistically identical and realistic but contain no real-world personal informationThe synthetic datasets allow clinicians, researchers, and commercial partners to safely complete the following tasks: **Train Machine Learning and AI:** Enables data scientists to develop and validate predictive models across multiple institutions without compromising patient confidentiality. **Run Cross-Border Collaboration:** Allows health systems and researchers globally to collaborate on early pilots and models without stepping into cross-border data privacy risks. **Test Platform Infrastructure: Stress- tests integrated care databases and analytics systems prior to launching live tools. Validate Research Findings:** Facilitates open science initiatives where researchers can share, replicate, and verify their findings by releasing "fake" datasets alongside their code.

The London Secure Data Environment (SDE) uses its artificially generated, privacy-safe **synthetic data** primarily to **train algorithms, test models, and conduct health research without ever exposing real patient records.**
As well as acting as an advanced analytics and innovation engine for the broader London Secure Data Environment (SDE), the London AIC links to the rest of the SDE ecosystem through several core functions:

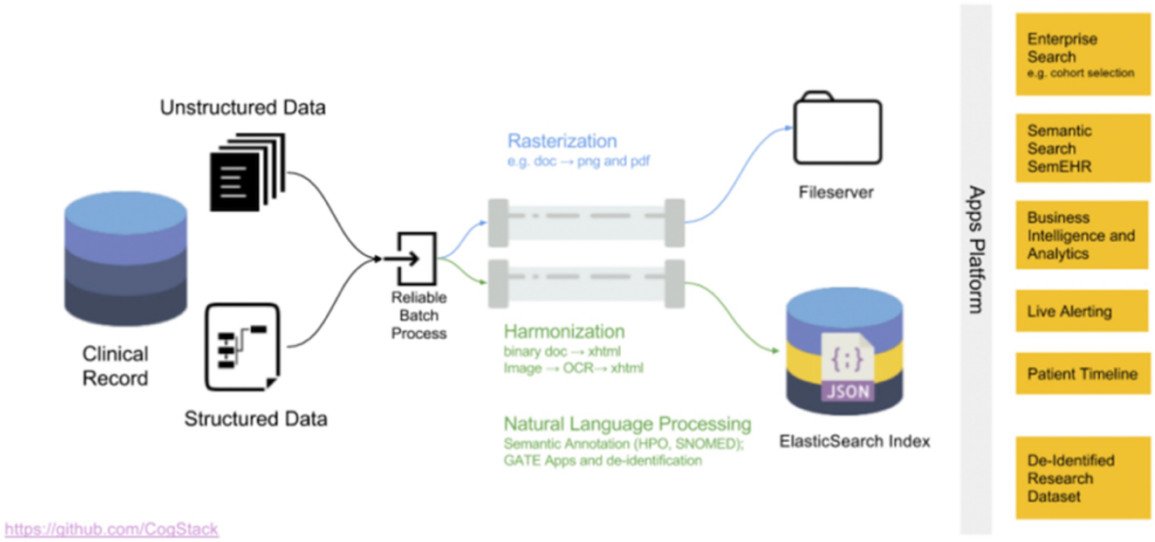
- **Platform Integration:** The AIC develops and deploys specialised platforms (like the Federated Learning and Interoperability Platform or FLIP) that plug directly into the broader SDE architecture covering London's **five Integrated Care Boards** (ICBs)
- **Advanced Analytics & AI Tools:** Rather than just storing data, the AIC builds and provides the artificial intelligence capabilities-such as predictive analytics and autonomous AI tools, and the clinical text extractor OncoLlama -that run securely on top of **the SDE's linked, de-identified datasets**.
- **Data Enrichment:** The AIC helps extend real-world data enrichment by integrating datasets from primary and secondary care (including complex image data and multi-modal records) and accelerates the validation of data for research from taking days in the past, to currently taking as little as 30 minutes. **This allows researchers and clinicians across London to apply machine learning more rapidly** to one of the largest health datasets in the world.

The role of NVIDIA: While not a financial investor in the traditional equity sense, Nvidia provides extensive in-kind investment by supplying computing resources, supercomputing clusters- such as access to Isambard-AI for the London AIC- and hardware infrastructure to train the multimodal medical data.

Data flow across the London SDE/AIC



Individual elements:
CogStack Infrastructure



Rasterisation is the process of converting mathematical shapes, lines, or 3D models into a grid of pixels (a 'raster' image) that a screen or printer can display.
Harmonisation is the process of extracting, standardising, and unifying disparate and unstructured healthcare records into a coherent, queryable format. It transforms fragmented Electronic Health Records (EHRs) into actionable, AI-ready datasets.
Elasticsearch is **the primary data store and analytics engine** used to process and fast-query unstructured Electronic Health Record (EHR) data. Processed EHR free-text and annotation documents (often extracted via Natural Language Processing/NLP) are stored as **JSON** (simple text-based format documents readable by humans and machines) in Elasticsearch indices

CogStack

Is used in primary and secondary care. It is an AI-powered healthcare information retrieval and extraction platform. On the Secure Data Environment (SDE) it serves as a search and analytical engine that automatically reads, structures, and unlocks vast amounts of unstructured patient health data to improve clinical care and medical research. Over 80% of NHS patient data is locked in un-structured free text e.g. letters, reports, and discharge summaries. CogStack extracts usable clinical meaning by mapping text to standardised 'medical terminologies' like SNOWMED-CT, and aids staff by accelerating the 'error-prone' process of assigning diagnostic and treatment codes at the point of care. It also allows researchers to identify potential subjects for studies quickly – taking seconds rather than months, and it runs real-time population health and caseload dashboards e.g. to flag up those 'lost to follow up', or on sub-optimal medication.
Live, frontline clinical environments, and complex, massive datasets, pulled from multiple sources can be 'surfaced' here (i.e. translated into simple, visually intuitive, metrics) on to dashboards.
Also, the generative AI model **Foresight** was developed by UCL and KCL researchers training it on de-identified health records on the National SDE facilitated through the British Heart Foundation Data Science Centre and the CogStack health informatics

platform directly alongside trusts in London. Please see details on[12] and international views expressed about Foresight (link). Foresight runs on AWS infrastructure

Data lakes

On the London SDE these serve as a foundational bedrock, securely ingesting and holding massive, unedited datasets (such as EHRs, medical imaging, and genomics) without needing upfront formatting. By keeping data in its full-fidelity raw format, researchers and the London AI Centre can train complex machine learning models, and authorised clinicians, academics, and **commercial researchers** can conduct population health analysis and develop proactive treatments. The lakes have three zones, the Raw zone which is the initial landing spot for all incoming clinical and demographic data, the Standardised Zone where data is de-identified and linked across different NHS Trusts, and the Curated Zone of highly refined data, which is made accessible within secure, accredited analytical workspaces.

AWS provides the necessary scalable cloud storage and computing resources, acting as a foundational environment.

- **Databricks** (in which all big tech have meaningful equity stakes): Provides the data intelligence and analytics platform layer, offering data pipeline capabilities and AI modelling advice.
- **Elastic** (all big tech have strategic partnerships with Elastic): Powers the underlying enterprise search[13]-allowing employees/users to search an organisations internal systems-and analytics engine functionality within the platform architecture

Sandpits

These function as:

- **Collaborative hubs** (physical or virtual): where clinicians, academics, and data analysts can query and link data sets, and investigate with pseudonymised patient data to find new ways to improve care pathways and inform health strategies, and
- **Research environments**: where approved researchers work on linked, secure population data to test new analytical models, algorithms and AI while remaining in full compliance with NHS privacy and security regulations.
- **Collaboration and proof-of- concept:** with the NHS, academia and **private enterprise** building prototype solutions or validating new technologies before deployment. They operate inside the secure infrastructure, ensuring that sensitive information remains anonymized or pseudonymised and subject to

[12] **Foresight** is an AI model specifically designed for healthcare that **has been trained in the NHS England Secure Data Environment (SDE)** on de-identified NHS data. Like ChatGPT, Foresight learns to predict what happens next based on previous events, working like an auto-complete function for medical timelines. Predictions are validated against real-world data. Foresight demonstrates significant potential benefits for healthcare delivery: **Population health insights**: Supporting better planning and resource allocation through analysis of data from diverse demographics, locations, and medical conditions, helping clinicians and policymakers anticipate health risks and plan interventions **Shifting from treatment to prevention** Predicting diseases early to enable timely interventions, supporting the NHS in moving towards more preventative approaches to healthcare **Precision medicine** Tailoring treatments effectively to improve outcomes and reduce side effects, potentially transforming patient experiences and healthcare outcomes **Reducing health inequalities** Ensuring inclusivity in predictive analytics by using data that represents the entire population.

[13] In the London Secure Data Environment (SDE), enterprise search typically refers to the use of the Health Data Research Gateway to explore and discover healthcare datasets across the UK NHS Research SDE Network. Researchers use this capability to find relevant clinical, demographic, and operational datasets

strict governance protocols governed by the pan-London Independent Information Access Group.

The Federated Learning and Interoperability Platform (FLIP)

A major output of technology development within the London AI Centre has been the Federated Learning and Interoperability Platform (FLIP). This provides an interoperable secure data environment within each hospital's firewall, a mechanism to query and analyse data across these environments ('Interoperability'), and a mechanism to train AI models ('Federated Learning') without the need to transfer or centralise data.

FLIP infrastructure overview

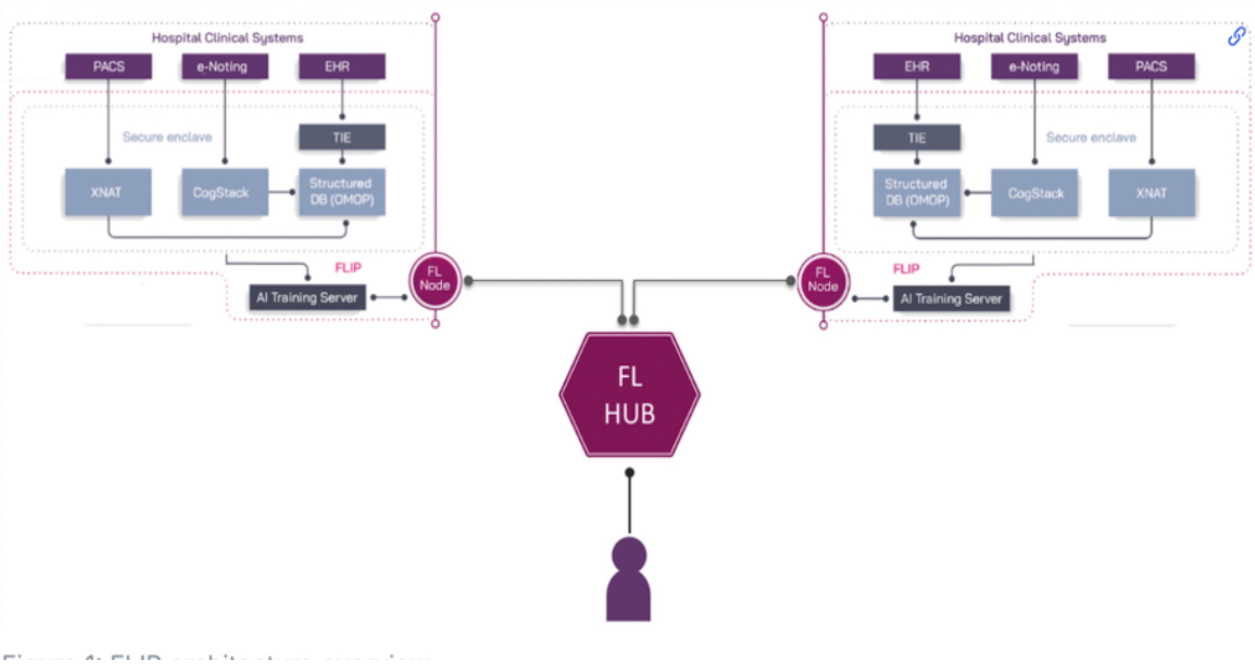


Figure 1.also see OMOP[14]

To populate in-hospital environments, data is loaded from multiple sources (Figure 1). These include electronic patient records systems (observations, diagnosis, medicines, etc) and data extracted from reports and narrative text through language AI systems (CogStack). These are transformed into OMOP. Information is also extracted from medical image stores within hospital Picture Archiving and Communications Systems (PACS), with additional capabilities for loading other imaging modalities and digital pathology. Imaging metadata is loaded into the OMOP radiology extension, enabling querying and selection of multi-modal data items.

The Radiology AI Summit 2026 were delighted to announce the launch of the Federated Learning and Interoperability Platform as a fully open source platform (Apache 2.0), by the team at the AIC with: strategic partner 'deepc' from Munich using AWS supported by OneLondon and Flower Labs[15] using AWS and with contributions from the NVIDIA

[14] OMOP is the Observational Medical Outcomes Partnership (Common Data Model) used to standardise diverse Electronic Health Record (EHR) systems- makes it easier to share, analyse and safely study clinical data.

[15] Flower Labs, is an AI startup specializing in federated learning and decentralized AI, uses **Google Cloud Platform (GCP)** for its cloud infrastructure, computing needs, and web-hosting operations. They maintain significant research partnerships with **Google** as well, frequently collaborating on developer groups and sharing open-source AI

NVFLARE team as the core federated learning engine working directly within the AWS environment.

Elastic, (see Data Lakes above) the core search, alerting, and information retrieval engine behind CogStack relies heavily on Elastic's infrastructure, with CogStack previously winning innovation awards. It uses AWS and Google to host it on the London SDE but can also use Microsoft Azure for infrastructure.

Figure 2: AIC on site infrastructure, currently installed in 7 NHS Trusts

> Each hospital hosts its own FLIP infrastructure, which includes Graphical Processing Unit compute (Figure 2) in an NVidia DGX-1 (8 x Nvidia A100). Processing is performed locally, eliminating the need to transfer vast quantities of imaging 'big data'.

Kubernetes (Cloud Providers): CogStack is designed to be highly flexible, running on Kubernetes clusters[16] that are deployed across AWS**,** and Microsoft Azure.

From a remote location, users can cross query information from multiple secondary care sites, perform analytics, and train AI models in a federated manner before bringing in model weights centrally.

The cluster of hospitals in which this hardware is likely to be available is those associated with a university; they obtain compute power from the Cambridge 1 supercomputer via their NVIDIA DGX cloud hardware.

---

advancements. In addition to Google, Flower Labs' developer stacks and ecosystems intersect with other major tech platforms such as AWS

[16] Kubernetes open source platform automates the deployment , scaling and management of applications, acting like a traffic controller and intelligently distributing network traffic. Scaling up and down according to usage.

To the author this all looks like a very expensive item of investment -even if only one or two hospitals in each of the 11 SDEs have such a platform. However, it is suggested that funding will come from:

- **National NHS Core Budgets:** NHS England used to provide foundational capital to build and maintain the pan-London SDE, distributing funds through multi-year government Spending Reviews. There is hope this will continue.
- **Integrated Care Board (ICB):** The five London ICSs are responsible for local digital transformation. They allocate portions of their regional budgets to ensure local hospital trusts can run FLIP enclaves within their secure firewalls.
- **Commercial and Academic Research:** The platform **IS DESIGNED TO HOUSE A DATA ASSET THAT CAN SECURELY GENERATE REVENUE**. Income generated through approved commercial licensing, health research partnerships, and NIHR-supported Driver Projects helps offset maintenance and operating costs.
- **Technology Partnerships:** FLIP is an open-source tool developed via a long-term collaboration between the AI Centre for Value-Based Healthcare (AIC) and its private-sector collaborators (incl. **deepc,** NVIDIA**, IBM, Siemens and GE Healthcare**) and which was built and deployed into production on AWS. The ongoing maintenance of it benefits from combined **public-private investments** and open-source scalability. Further evidence that **big tech** firms are integrated across the board as multinational industry partners 'to help scale real-world data usage and AI models within the NHS.'

Guy's and St Thomas' Hospital (GSST) as an example (also see diagram below)

has....

1. Connections to more than a dozen legacy electronic systems covering 20 years of hospital activity, and a live connection to the hospital Electronic Health Record, enable raw data to be ingested into a on-premise database server. This includes diagnoses, attendances, laboratory tests, clinical observations, medication prescribing and dispensing, as well as data from specialty systems, for more than 5 million unique patients. Within the database, dbt is used to standardise the data into a common data model, and then into OMOP. An analytics cluster allows deployment of dashboards and applications for internal use.

then....

2. Unstructured data from the live EHR and from 20 years of historical documents, letters, clinical notes, summaries, and test/radiology reports, is ingested into the CogStack server. Here, language AI models are trained and deployed to extract concepts with high precision into the structured data models.

and finally...

3. Standardised data is pushed into the secure FLIP enclave, where radiology metadata is used to pull DICOM from a live imaging data lake through XNAT, joined to the rest of the patient health record.

Guy's and St. Thomas' Hospital (GSTT)

GSTT is one of the 'home' Trusts for the AI Centre (Figure 1), and is used here to illustrate the type of work that is being conducted:

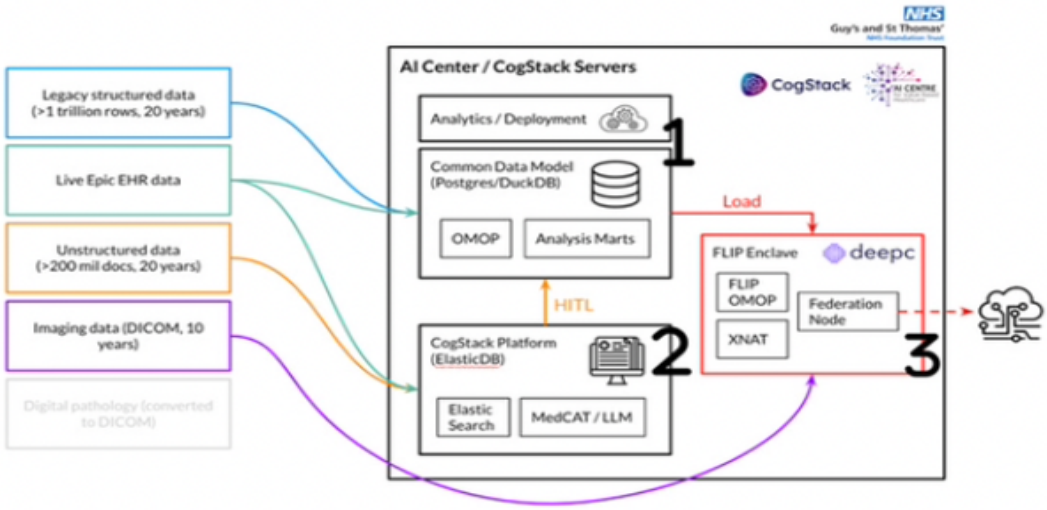


Figure 1: GSTT Real-World Data Architecture

dbt: (data building tool) sits 'natively' within the Microsoft cloud environment and is provisioned alongside other standard software tools like Python, Rstudio and Azure Machine Learning

DICOM is a machine interoperable Digital Imaging and Communication file for radiology. It bundles the image's raw visual data with a "header" containing vital patient information (e.g., name, birthdate, study parameters) in a single file. Microsoft provides Fabric as a data transformation capability and Azure Health Data for hosting services.

XNAT (eXtensible Neuroimaging Archive Toolkit) developed at Washington University imports, archives, processes, and securely shares medical imaging data (such as MRI, CT, and PET scans) alongside associated clinical records for research purposes. While electronic health record (EHR) systems like Epic (on Microsoft) or Cerner (on Oracle) store standard patient history, medications, and clinical notes, XNAT acts as a specialized, research-focused adjunct that can strip Protected Health Information (PHI) from DICOM image metadata, blur faces, and remove burn-in data, turning raw clinical patient scans into anonymized records for clinical trials.

Hospitals use XNAT to securely transfer specific patient imaging studies from hospital PACS (Picture Archiving and Communication System) servers to research nodes where they are used to evaluate or train Artificial Intelligence (AI) diagnostic models and conduct retrospective studies.

XNAT also features a data model that allows researchers to attach clinical and genomic data to the same subject. This means researchers can cross-reference patient demographics or biomarker data against their imaging results. It is also widely used to

centralise patient data across multiple hospitals for large-scale or multi-site medical studies.
So, while a clinical PACS is built for immediate, daily clinical diagnosis and patient care, XNAT is built for data exploration, algorithm testing, and scientific collaboration. While it is big tech agnostic, XNAT has been integrated into Microsoft on the London SDE.

Data science teams for ICBs: the AIC is providing practical support in clinical informatics, data science, and machine learning (ML) development and deployment. Primary aims are to: (a) help integrated care boards (ICBs) migrate data pipelines[17] and analytics into common data models and terminologies within LDS environments; (b) extend these into reproducible pipelines for data science, predictive analytics deployment and MLOps[18]; and (c) work together and provide training to make ICBs self-sufficient in these capabilities. The AIC will also support the adoption and roll out of the OMOP Common Data Model to enable future linkage to hospital data environments.

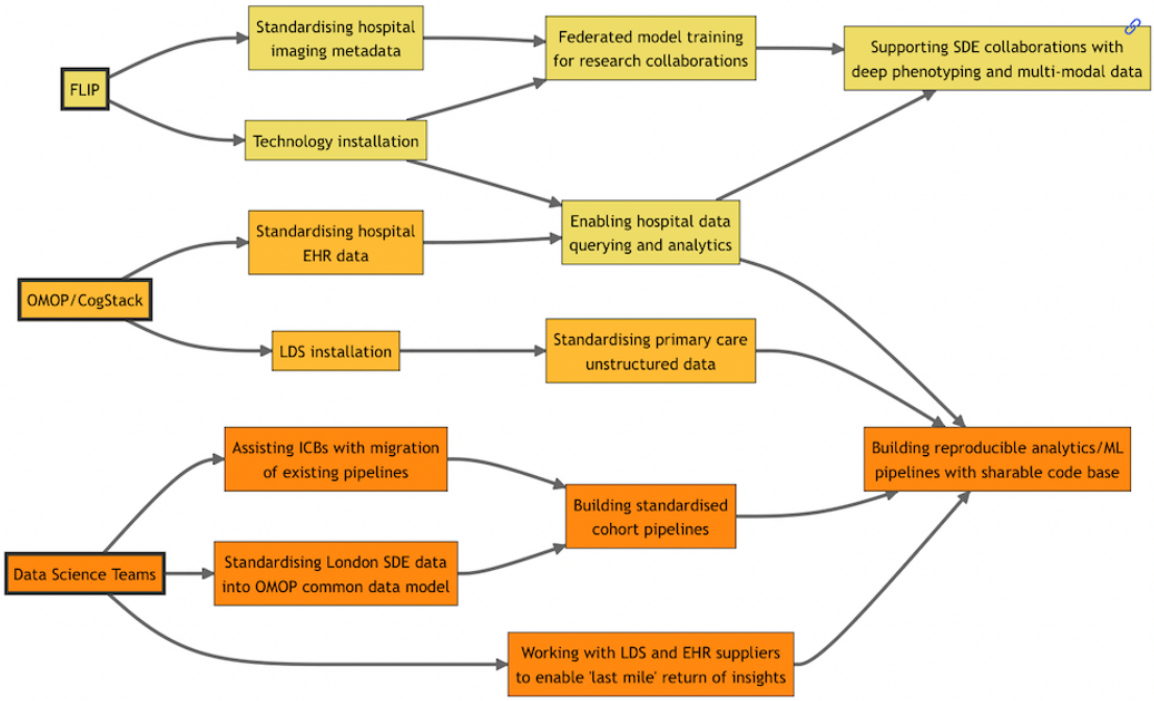


Figure 1: Summary of AI Centre work components and objectives. FLIP = Federated Learning and Interoperability Platform; ML = Machine Learning; EHR = Electronic Health Record

Roving Teams

The London AIC operates through 'Roving Teams' which include those collaborating with NHS hospitals and population health centres such as ICBs. Teams may include

[17] A data pipeline is a set of automated processes that moves raw data from multiple sources to a destination (like a data warehouse or data lake). Along the way, it cleans, filters, and transforms the data so it can be safely stored and used for business intelligence, analytics, or machine learning.

[18] a set of practices that combines machine learning, data engineering, and DevOps to automate and standardize the entire lifecycle of AI models. Its core goal is to bridge the gap between model development and operational deployment. DevOps itself is a working culture and set of practices that unifies software development (Dev) and IT operations (Ops). By relying on automation and continuous feedback, it eliminates traditional team silos to deliver applications and services at a much higher velocity and reliability.

engineers, data scientists, and developers who are working to install technology, and build data pipelines and analytics products. Team members are employees of NHS Trusts and Universities affiliated with the AI Centre and may hold honorary contracts to enable work at other locations.
Team members and partner engineers hold advanced industry certifications (e.g. AWS Certified Solutions Architect, Google Cloud Professional, Microsoft Azure Solutions Architect). The NHS delivers these capabilities by partnering with accredited IT vendors and strategic cloud partners like Nordcloud and Telefónica Tech[19], who possess the specific qualifications required to connect legacy trusts to hyperscale cloud environments. All cloud linkages are governed by strict national mandates and zero-trust security platforms, such as the NHS Secure Boundary service provided by IBM and Palo Alto Networks. These roving and engineering teams follow rigorous architectural roadmaps detailed in the NHS Cloud Journey Guide, ensuring that connecting trusts to public and private clouds maintains strict 'data sovereignty' and compliance.[20]

London Data Service (LDS)

Data Extraction & Linking:

The North East London Integrated Care Board (NEL ICB) is responsible for developing and hosting the LDS. This involves:
Data collection: secure collection of patient information from primary care (General Practices), acute hospitals, and to and from NHS England via the Federated Data Platform (Palantir Foundry on AWS and Microsoft infrastructure) for national databases/comparisons etc – this includes for 'faster data flows'.
Record Unification: merging information collected to create comprehensive, longitudinal health records across all care settings.
Data Preparation: organising this data into identifiable forms (for direct patient care) and de-identified/pseudonymised forms (for academic and commercial research).
Platform Provisioning: provision of carefully prepared datasets into secure analytics workspaces (such as the London Analytics Platform).
The London SDE aims to create and maintain one of the largest and most comprehensive health data assets in the world, covering London's population of over 10 million residents. The LDS makes this possible by streamlining how data is ingested and standardised.[21]

Common Data Model/Digital Image and COM

Bringing data from multiple systems and applications together is an expensive and time-consuming task because when you are not able to share and understand the same

[19] Nordcloud and Telephónica Tech are both subordinate customer service 'partners' of big tech with the former holding top-tier audited certifications with Microsoft, AWS and Google ecosystems, while the latter has a 'strategic' partnership.
[20] In the NHS, **data sovereignty** simply means that patient health information remains entirely within UK jurisdiction, and it is legally protected from foreign laws and external access. **Compliance** ensures that this data is processed, stored, and shared according to strict UK legal and ethical frameworks, most notably the UK General Data Protection Regulation (UK GDPR). It is NOT a state/people owned and run data stack that many campaigners wish for.
[21] You can learn more about how the LDS manages patient privacy and information on the OneLondon SDE Privacy Notice https://www.onelondon.online/wp-content/uploads/2025/11/London-Data-Service-Privacy-Notice-November-2025.pdf or explore the broader ecosystem at the OneLondon London Secure Data Environment https://onelondon.online/london-secure-data-environment/ website.

data easily, each application or data integration project requires a customised implementation.
'Common Data Model' simplifies this process by using a 'metadata system'[22] to provide a shared data language for different applications to use. It includes a set of standardised, extensible data **schemas** that Microsoft and its partners have published. More information can be found on the Common Data Model repository on GitHub (GitHub is owned by Microsoft).
To summarise the benefits:

- Structural and semantic consistency across applications and deployments.
- Simplified integration and disambiguation of data that's collected from processes, digital interactions, product telemetry, people interactions, and so on.
- A unified shape, where data integrations can combine existing enterprise data[23] with other sources and use that data holistically to develop apps or derive insights.
- The ability to extend the schema and Common Data Model standard entities to tailor the model to your organisation.

The **DICOM** (Digital Imaging and Communications in Medicine) standard unites both **image data** (the visual scan) and rich **metadata** (the patient's data, scanner settings, and demographics) in a single file. **COM** typically refers to **DICOM Communication** operations, specifically the **C-STORE** protocol used to transmit these images across networks.

THE LONDON SDE AS AN 'AI FACTORY'

- It is 'a specialised infrastructure designed to continuously turn raw data into intelligence, serving as a modern "power station" that produces digital goods rather than electricity.'
- In the context of technology and healthcare computing, it is 'an industrialised, end-to-end platform engineered to automate and scale the entire artificial intelligence lifecycle. Instead of simply generating data, the London ecosystem can use continuous data ingestion to rapidly train, fine-tune, and deploy predictive intelligence'.
- It possesses the three necessary core components for an AI factory within the local tech and health sectors**:**
    1. **HIGH-PERFORMANCE INFRASTRUCTURE:** as provided majorly by Microsoft and AWS, with access to massive, specialised compute power (see Isambard (NVIDIA) in Bristol and Dawn (Intel Ponte Vecchio) in Cambridge, both super computers configured for extreme speeds and parallel processing, and **full-stack integration** with physical hardware, pre-trained foundation models, and scalable deployment pipelines (see below) in a single operating environment.

[22] Metadata is simply defined as "data about data." It provides descriptive, contextual, or structural information about a specific file, document, or dataset. Instead of being the content itself, it explains what the content is, making it easier to organize, search, and manage.
[23] all the digital information any 'company' gathers, stores, and uses to power its daily operations and long-term strategy. It forms the foundation for 'business intelligence', encompassing both structured information (like databases and spreadsheets) and unstructured content (like emails, video, and IoT sensor logs)

2. **ACCESS TO A SECURE DATA ENVIRONMENT (SDE):** In the UK's capital, AI factories are closely tied to healthcare and life sciences, integrating fully with the **London SDE** with all their data.
3. **THE "DATA FLYWHEEL" (AKA CONTINUOUS IMPROVEMENT)** Like an AI factory the London SDE is not just a static data centre. It functions 'like an assembly line' with:

Ingestion: Raw, multimodal data is automatically cleaned and processed
Training & Fine-tuning: Pre-existing AI foundation models are tailored for specific domain applications (e.g., medical diagnostics, predictive healthcare)
Inference: Models are rapidly deployed into live production
Automated Monitoring: The system continuously evaluates the model's outputs in real time, capturing feedback to make the next iteration of the model faster and more accurate

The pipelines necessary to take research, AI, and health-tech products to market in the London **SDE**, are collaboratively supplied by OneLondon-a partnership of London's five Integrated Care Systems (ICSs)- and health innovation networks. Of major note these include infrastructure suppliers for the entire system Microsoft, AWS and Google.
The pipelines of:
The regional data platforms and federated analytics infrastructures such as those produced by the **Artificial Intelligence Centre for Value-Based Healthcare (AIC)** and NHS partnerships built on AWS.
The London Data Service (LDS) hosted in North-East London, which acts as the data engineering layer that extracts and links pseudonymized records across primary and secondary care and is built on Microsoft.
The London Analytics Platform built on the previous *Discover-NOW* model, acting as the primary pipeline allowing academic and commercial industry partners to develop, train, and validate health-tech, and built on Microsoft.
The London AI Centre (AIC) as a hub that supplies frontier technology and expert teams, enabling developers to test algorithms on real-world data and multi-modal datasets built on Microsoft.
Local Secure Pipelines: individual health nodes, such as the Paddington Life Sciences (Microsoft and Google) initiative and the iCARE SDE (Microsoft and Google) supply localised real-world health data across major hospital groups such as Imperial College Healthcare.

To utilize these pipelines to bring a product to market, you typically need to undergo secure data accreditation, with data access criteria and requirements on the NHS England Secure Data Environment portal. These will depend upon

- What **specific type of product** you are looking to take to market (e.g., AI diagnostic tool, patient pathway optimization, therapeutic drug evaluation).
- Whether your organization has already achieved Data Security and Protection Toolkit (DSPT) accreditation.

Big tech firms operating in London possess the industry's most rigorous secure data accreditations. For specific sectors—such as healthcare these providers meet exceptionally strict public sector compliance standards. They are ready and waiting!

The marketable outputs of an AI factory are often called "tokens" and represent a new form of digital capital, including predictive insights, automated decisions, and generative content. The key marketable outputs of an AI factory are:

AI Agents and Autonomous Systems which are considered the highest-order output, acting as digital workers that can plan and execute tasks. Examples relevant to the NHS are

1. Autonomous Agents Software: that can take actions to achieve goals, such as managing supply chains or controlling robotic systems.
2. AI-Enabled Products: Vertical AI solutions that package models into workflows for sectors like healthcare.
3. Actionable Intelligence and Predictive Insights outputs to help organisations anticipate trends and make data-driven decisions.
4. Predictive Analytics: Forecasting patient need, unplanned/urgent attendances, or equipment failures to optimise operations.
5. Risk Analysis & 'Fraud' Detection in the NHS
6. Drug Discovery & Scientific Discovery: Accelerating research by analysing complex biological datasets.

Generative AI Content and Tokens. AI factories produce vast quantities of "tokens" used in generative AI applications.
Content:

1. Text and Language Models: Generating code, marketing copy, or technical documentation via Large Language Models (LLMs).
2. Visual Data: Creating or interpreting images (and videos) for radiology examinations or patient interviews for diagnosis, symptom development and training etc.
3. Process Automation Solutions: these solutions reduce operational costs by automating routine tasks.
4. Robotic Process Automation (RPA): Automating office tasks such as invoice processing, patient 'customer' service, or data entry.
5. Computer Vision Solutions: Utilising visual data for automated quality control in manufacturing or monitoring safety in public spaces.
6. Customised Models and Fine-Tuning Services. AI factories allow for the creation of specialised, discipline-specific AI models.
7. Domain Adaptation: Fine-tuning base models on private, enterprise-specific data to create proprietary intellectual property.
8. API Services[24]: Offering specialised AI functionalities for developers to integrate into their own applications.

By converting data into these actionable outputs, AI factories allow companies/organisations to generate revenue by selling intelligence and efficiency at scale.

[24] API stands for Application Programming Interface. In the context of APIs, the word Application refers to any software with a distinct function. Interface can be thought of as a contract of service between two applications. This contract defines how the two communicate with each other using requests and responses.

Tokens: Tokens are important and valuable in health data because they convert sensitive personally identifiable information (PII) into non-sensitive, unique identifiers (tokens).
This process, known as tokenisation, ensures privacy compliance (such as HIPAA and GDPR) while allowing researchers and providers to connect disparate health records, track patient journeys over time, and secure data against breaches.

Here is why tokens are important and valuable in health data:

1. **Privacy Protection and Security**
   Irreversible Security: Unlike encryption, tokenised data is typically non-reversible without access to a secure "token vault". If breached, the tokens are meaningless to unauthorised users.
   Compliance: Tokenisation helps organisations adhere to strict regulations (HIPAA, GDPR) by ensuring that actual patient identifying information is not exposed in analytics.
   Reduced Liability: By replacing sensitive data with tokens, organisations reduce the risk and cost associated with data breaches.
2. **Linking Disparate Datasets (Interoperability)**
   Tokens enable the linking of patient records from different sources—such as claims data, laboratory results, and medical records—without exposing the patient's identity.
   They allow researchers to track patients over time and across different settings (e.g., from trial participation to real-world clinical usage).
   Researchers can use tokenized data to accelerate clinical trials, improve study designs, and gain deeper insights into drug safety and efficacy.
3. **Enabling Data Utility**
   Tokens are crucial for Natural Language Processing (NLP) and AI applications, allowing models to process and analyse data securely.
   Tokenisation helps pharmaceutical companies analyse patient behaviour and treatment effectiveness in real-world settings.
   By enabling the analysis of large, aggregated datasets, tokenisation supports more personalised treatment approaches.
4. **Patient Empowerment and Incentives**[25]
   Tokenised incentives—often using blockchain or Non-Fungible Tokens (NFTs)—reward patients with digital tokens for sharing their data or achieving health milestones (e.g., logging steps or taking medication).
   NFTs are also being explored to give patients control over who can access their health data, improving transparency and enabling them to monetise their personal information.
5. **Health and Wellness:** Patients can also be rewarded with utility tokens for meeting wellness goals (e.g., hitting step counts, completing health surveys, or

[25] **Fungible tokens** like utility tokens or healthcare cryptocurrencies are identical, interchangeable, and easily divisible, making them ideal for patient reward systems, tokenomics in telemedicine, and decentralised health data economies. **Non-fungible tokens** (NFTs) are unique, indivisible digital certificates that authenticate a single, specific asset. In healthcare, NFTs are heavily utilised for patient-centric electronic health records (EHR) management, clinical trial consent, supply chain tracing, and tracking organs for transplant

adhering to medication schedules). These tokens can then be redeemed for health-related products or discounts

6. **Patient-Owned Health Records (EHR):** Patients can "mint" their health data as an NFT, effectively owning it. The NFT acts as a secure digital key or access log, allowing patients to grant, track, and revoke data access to specific doctors, specialists, or researchers.
   Patients can securely lease their medical history (such as anonymised genetic or wearable device data) to pharmaceutical companies for research, earning royalties or payments each time the data is accessed.

Tracking life-saving drugs from the manufacturer to the patient. NFTs represent unique batches or individual crates, ensuring medications are authentic and preventing counterfeit drugs from entering the supply chain.

In summary, tokens are meant to unlock the 'value' of healthcare data, transforming it from siloed, sensitive information into a secure, interoperable resource that drives better research, improved patient outcomes, and greater efficiency in the healthcare system.
Also managing digital consent forms for clinical trials and keeping tamper-proof records of patient credentials; medical degrees; and diagnostic imaging.

Part two:

Concern for workers with government's demands for increased 'productivity' and an agenda focused on digitalisation.

The new digital NHS healthcare services and analytical/healthcare tools are increasingly being developed and deployed as the NHS becomes increasingly based around the 11 regional and one National Secure Data Environments (SDEs)- big tech 'cloud' based infrastructures for generating tech 'solutions' and commodifying our NHS data in a value-based way[26].

As of 2026, the NHS is attempting to resist complaints about the lack of safety for our data by shifting from a model of sharing data to one of accessing data, with these SDEs providing the secure, **regionalised** platforms for this new paradigm, and the National SDE acting as a central data 'safe haven' and holder of the '**single source of truth**' for the entire NHS system. This is enabling greater use of data driven services across the NHS.

Key Aspects of the NHS SDE Strategy (2026):
Regional Structure: The 11 sub-national SDEs, which cover between 5-12 million people each, are being established to manage data regionally, focusing on deep, multi-modal data such as imaging, genomics, and local datasets.

[26] Using human data to improve health care efficiency by reducing cost of services (fewer / briefer/ lesser or skill-mixed staffing); attempt to prevent need for services; use services in a value-based way (deny aspects of care); and use digital as a part of or full alternative for human deployment.

A network of 'Data Safe Havens': The SDEs act as "data safe havens" where data remains in a secure environment while authorised researchers and analysts work on it, without taking the data out. Their focus is both on research and the production of *new treatments, clinical tools, and predictive models*, such as the SLAIDER[27] project for respiratory admission prioritisation, as well as improved *clinical planning and resource allocation*, all the while integrating with any current National Strategy. It has been important to ensure consistency and interoperability across the country.

Digital Transformation: The moves to shift the NHS from 'analogue first' to 'digital first' and from 'hospital' to 'community' will require *a newly trained workforce* with new digital ways of working. The use of remote monitoring, and *digitally enabled personalised care* are already well underway, and new, larger digital structures such as *'virtual wards'* seem to be announced nearly every week in med tech magazines.

SDEs and Cloud Technology: The SDEs are all big tech cloud-based and provide the patient related information and analytics clinical staff need for their geographical areas. Whilst many NHS providers may have made their own/different choices from the listed NHS approved software for various tasks such as for triage and assessments, the London SDE significantly improves interoperability by unifying and linking all patient records and social care records across the ICSs. The London SDE has its own 'foundational data warehouse' using Snowflake software on Microsoft infrastructure that can feed all of London's points of care.

Examples of developing staff technology: Secure workspaces often include a *virtual desktop* and for general practices often EMIS (AWS) or TPP (heavily integrated with Microsoft and AWS) software, while hospitals use *Epic* (on Microsoft) or *Oracle Health*- which has had a long standing partnership and pooling of cloud technology with Microsoft.

Total GP triage systems are increasingly popular in London with clinical review of submitted requests determining response such as in-person, phone/video, self-care or referral to another service. Examples are *Accurx* (Microsoft), *eConsult*, *PATCHS, Klinik Healthcare*, *Anima and Rapid Health* (all AWS).

Nurses in Same Day Emergency Care (SDEC) units which are usually heavily aligned to 'virtual wards' use *EPR systems, digital clinical assessment software, and mobile devices* to rapidly assess, diagnose and treat on the same day. The SDEs also enable '*patients' longitudinal linked records'* to be analysed, an aspect critical for the 10-year health plan's digital goals. It should be noted that the development of SDEs will mark a **significant consolidation** (i.e. job cuts) **of regional capacity**, in a sense they are the new structure of the NHS.

Whence productivity

Any increased productivity will not be based on investment in more skilled staff as required by unions, but by moving as much assessment, treatment and ongoing care into the community as possible, and having AI assist staff functions to improve

[27] https://le.ac.uk/computing-and-mathematical-sciences/research/life-health-sciences/slaider

‘productivity’ by: individual nurses seeing more patients in the time given; lower nursing grades (see below) being enabled by AI and supervision to perform more complex tasks; reducing the need for community services to refer patients to other more expensive services such as outpatients, Urgent Care units and Emergency Units.

The most significant AI tools being deployed across London currently include:

Automated Scheduling and Routing

Community nurses spend significant time coordinating schedules and traveling between home visits. AI engines (such as those piloted by Care City CIC with NHS ICBs and Xantura (aquired by Newton in 2024 and partner of Microsoft through its partner Oneview and Google through Feebris) analyse patient needs and locations to automate complex daily rosters. Coordination work that used to take hours is reduced to minutes, minimising travel time and ensuring the right nurse is matched to the right patient

AI-Assisted Wound Care

Trusts, including the West London NHS Trust, utilise smartphone technologies like *Minuteful for Wound* (Apple and AWS) to photograph and precisely measure chronic wounds. Measurements and healing progress are instantly uploaded to a secure portal, with some trusts noting that consultations are over 5 minutes faster and enable more accurate, consistent care tracking.

Proactive Patient Monitoring:

Across London, health networks use predictive AI algorithms to analyse clinical data and screen patients for early signs of deterioration. For instance, programs like *Minder Care* (Microsoft Kubernetes cluster via Imperial) use sensors and AI to monitor at-home dementia patients, alerting community teams to risks like falls or night time wandering. Nurses and clinical coaches can intervene before an emergency hospitalisation occurs.

Administrative Time-Savers:

With the rollout of productivity tools like Microsoft *365 Copilot* across what was NHS England, community staff can use AI to instantly draft care plans, transcribe visit notes, and summarise patient records.

Staff save an average of up to 2 days per month on paperwork, leaving more time for actual clinical duties.

Remote home care/Virtual wards:

These will require **a Kit & Monitoring Devices** which patients are provided with to take home. Essential hardware typically includes:

Pulse Oximeter: To measure blood oxygen levels and pulse.

Blood Pressure Monitor: For routine cardiovascular spot-checking.

Thermometer: For tracking body temperature.

Wearables: Continuous monitoring devices worn on the arm to track heart rate, breathing rate, and temperature.

Specialised Devices: Scales for heart failure patients, and portable devices like bladder scanners or point-of-care ultrasound (POCUS) brought in by visiting community nurses.

Communication and Digital Inclusion Tools: Patients require an easy way to communicate with the virtual ward hub without relying on any personal tech.

Tablets / Digital Hubs: Pre-configured tablets that walk patients through their monitoring schedule and allow for video consultations.

4G / Wi-Fi Hubs: Secure local networks provided by the NHS so that data transmission does not rely on the patient's home internet connection.

Central Clinical & Software Platforms for NHS staff

e.g., at Kingston Hospital or South West London ICS, the back-end requires**:**

- **Remote Monitoring Dashboards:** Software that securely aggregates incoming vital signs and alerts specialists if a patient's condition deteriorates.
- **Interoperable EPR/EMR Systems:** Systems that seamlessly integrate with standard NHS software networks (like EMIS Web (AWS), SystmOne (AWS/Microsoft), or NHS Spine (AWS) to ensure care teams can access full medical histories.

Data Security Compliance: All technology must meet strict national regulatory frameworks, including the Digital Technology Assessment Criteria (DTAC). [see link about virtual wards1]

Private providers of virtual ward structure

London does not have a single provider for its virtual ward structure. Instead, the NHS in London works with several healthcare technology and clinical service companies to deliver "Hospital at Home" models. The primary providers currently operating across the capital include

Doccla (Google)**:** One of the largest virtual ward providers in the UK, partnering with NHS trusts across London (including Royal Free London[28] and South East London ICS) to provide remote patient monitoring kits and health coaching

The Access Group (Microsoft)**:** Supplies digital integration and electronic patient records software to manage virtual ward operations for trusts like NELFT. [1, 2, 3]

HomeLink Healthcare (through Symbiant and Mobizio-now 'Access Care Planning' (AWS): Provides an all-in-one service combining remote monitoring equipment with community-based clinical teams and 24/7 on-call support

Spirit Health (Microsoft): Collaborates with NHS trusts like Whittington Health to provide virtual ward pathways for vulnerable Londoners.

The Creation of "Associate" and "Enhanced" Roles

The NHS Long Term Workforce Plan aimed to expand intermediate, non-traditional roles with an initial rise from 1% to 5% of the total workforce using technological integration examples include: **Nursing Associate (NA) and Physician Associate/Assistant (PA)** roles both of which are causing a lot of concern from Professional bodies. The NA is intended to bridge the gap between Health Care Assistants (HCAs) and Registered Nurses (RNs). Digital care plans and structured software tools are proposed to allow them to administer certain medications and manage treatments with clear, app-guided compliance protocols.

The trained PA will use digital diagnostic paths and triage software to manage standard patient lists under the supervision of a Senior GP or Consultant.

---

[28] The ex CEO of RFH became a member of the DOCCLA board after his retirement. He was also involved in the transfer of 1.6 million identifiable healthcare records from RFH to Google Deepmind without a contract. https://www.digitalhealth.net/2023/09/sir-david-sloman-departs-nhs-england-to-join-doccla-as-senior-advisor/ https://www.healthcaredive.com/news/uk-watchdog-google-deepmind-nhs-out-of-step-with-patient-privacy-laws/446313/

The Core Boundary: Regulation and Safety
Health think tanks emphasise that digital tools do not eliminate the legal necessity for human judgment. Under NHS Workforce Safeguards, any task shifted to a lower-skilled worker via technology must be accompanied by explicit delegation protocols, evidence-based toolkits, and clear lines of clinical accountability.

However, in the Tech Sector: entry-Level to Specialist Pathways
While clinical progression is currently locked behind strict medical degrees, the **Digital, Data, and Technology (DDAT)** sector within the NHS will allow entry-level of non-clinical workers to step quickly into skilled technical positions. Workforce, training and education will be through:

- **Data Academies.** The NHS Digital Academy provides rapid upskilling pathways for staff members who previously worked in basic clerical, portering, or domestic roles to retrain in data management, software support, or cyber security.
- **'Digital Champions'** will be existing non-technical staff actively upskilled through specialist training courses to help their clinical peers to navigate changes such as new EPR rollouts, software updates, and hardware interfaces.

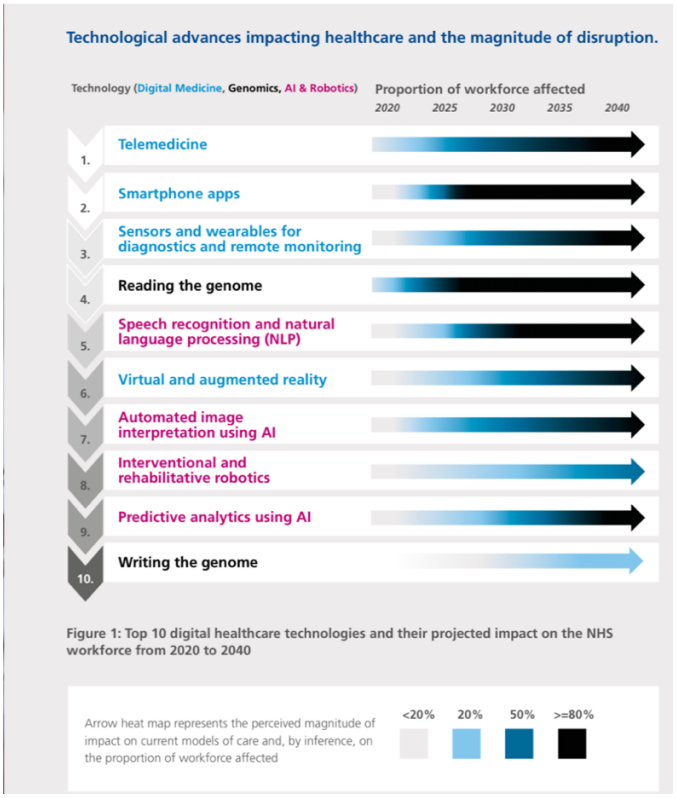


Figure 1: Top 10 digital healthcare technologies and their projected impact on the NHS workforce from 2020 to 2040

The above chart is from The **Topol Review**, published in 2019 by Health Education England and led by Dr. Eric Topol, is a landmark strategic report outlining how to prepare the NHS workforce to deliver a digitally driven future. It details the education and training required for clinical staff to integrate genomics, digital medicine, artificial intelligence, and robotics into daily healthcare

The use of the Secure Data Environment for financial gain/national productivity:
Simplistically-but of some help to clarify thoughts-within the current paradigm of capitalism, increased 'productivity' and/or profit can only be achieved by either spending less money on means of production, be it human labour and/or 'machinery'

and other components, or by achieving more profit in the market without the need for any reduction in production spend. We are currently in a state of chronic global economic stagnation and present politics mean that public service spending is under constant threat of reduction.
However, the global economic slow-down began before the big tech companies (like Microsoft, Google, AWS, Oracle and Nvidia) had realised the value of large quantities of human data in training and improving AI (a key aspect of digital capitalism) for use in public ecosystems, and the income available from renting out the products from the commodification of that data. They also found that when working in Research Partnerships, their early access to healthcare datasets allowed them to collaborate on machine learning models on their cloud, giving them huge leverage over future health-tech patents and intellectual property (intangible assets). With public researchers this has historically meant big tech taking over 99.7% of all IPRs generated by joint research.

The way ***the London SDE data environment has been set up it should be seen as a utility*** and London NHS worker power, as well as providing a service, is providing valorisation of corporate investment.

The routes to market:
Big tech have monopoly powers in many of the 'routes to market' in the healthcare digital economy, and both the billion dollar venture capitalists who invest in big tech corporations, and big tech themselves have invested billions of dollars in our NHS while assuming critical infrastructure positions throughout our 'digital utility'. It is likely that the other SDEs will follow suit. This puts big tech into powerful positions to dominate in any negotiations around the division of financial returns, and such a profit stream could also become longstanding as the 12 SDE systems begin to work and become 'clinically essential'. Over time it may become financially challenging for the UK to remove itself from big tech presence (ecosystem 'lock-in') and as well as generating profits big tech will have managed to 'take a seat' in our health care as an essential partner to government.

END

.