techUK
FOR WHAT COMES NEXT

How can the technology industry help tackle challenges around Secure Data Environments?

Background

The government is investing in the NHS Secure Data Environment (SDE) Network, established to centrally coordinate the delivery of a small suite of interoperable SDEs, unified through a Community of Practice to ensure the adoption of consistent and interoperable standards and technology. This Network is designed to accelerate the discovery of new treatments and help the NHS to plan better and more effective services.

techUK, the UK's technology trade association, has welcomed this initiative from the outset. The techUK Life Sciences Working Group (LSWG), bringing together members actively working in drug discovery, digital therapeutics, data, and AI, aims to help identify how the technology industry can facilitate and deliver the ambition to set the UK apart as a world leader in life sciences and deliver life-changing innovations to patients. Both the digital health and life sciences sectors see SDEs as a critical opportunity, poised to take advantage of the world's largest linked datasets and make the UK the best place in the world to develop innovations that can help the NHS.

This document puts forward proposals for some of the challenges facing Secure Data Environments over the next 6 – 18 months, on which SDEs and industry could work together. SDEs pose a fantastic opportunity for the UK, but it is critical that we move quickly and ensure the underlying architecture is fit for purpose.

Challenge 1: Technical consistency and maturity

SDEs across England are at different stages of maturity and have varying local infrastructure preferences. As a result, it is likely that SDEs will adopt a variety of architectures, creating a patchwork data ecosystem which will hinder interoperability between SDEs and their overall functional utility. The SDE Network's Community of Practice is intended to prevent the uncontrolled proliferation of otherwise 'siloed' and small scale local SDEs that cannot federate, with the NHS Research SDE Network intended as the primary route to access data, as set out in the NHS' Data Access Policy.

Given that improved data quality, data mapping, research feasibility, and cohort identification will be rapidly required at scale, this may cause challenges. For example, a pharmaceutical firm's R&D team looking to establish the best locations for a new research site for a Phase III study will want a rapid assessment of patient population numbers in different SDEs – this would require consistent searches to be run across comparable data sets to give analogous outputs. In addition, ensuring that the research, life science and digital health industries can link their own datasets with NHS datasets held in SDEs, will be critical to maximising the value of these environments.

The NHS Research SDE network and SDE owners should work with industry across a number of areas to tackle these challenges:

- The SDE Network plans to surface metadata and a service catalogue on the Health Data Gateway SDE collection, as well as other data assets across the UK. **Industry would benefit from early visibility of technical standards, common data models and data completeness and quality**, across SDEs including highlighting opportunities where suppliers can support the SDEs and/or the NHS Research SDE Network to address specific challenges.
- Industry would welcome the setting out of **common policies and technical standards across SDEs to ensure interoperability**, ease of connectivity to external industry datasets, and consistent use of data standards to derive common outputs (e.g. feasibility counts) across regions.
    - While the SDE Network's Tech and Data Working Group is planning to manage this within a Community of Practice, it is critical to ensure both life science and technology industry representation within this group to optimise outcomes, and to return actionable knowledge back to front line services.
- Involve industry and other provider organisations to ensure **consistent data linkage processes and common ontologies for other key datasets**, including those currently under-represented in research such as mental health and social care; and other Social Determinants of Health (SDOH) data.
- **Test the ability to federate** through a series of seven federated driver projects planned to be up and running by March 2025. These projects should be an opportunity area for industry to support the SDE networks.
- **Progress the adoption of OMOP to support interoperability,** including enabling cohort discovery, and the development of Bring Your Own Data functionality as part of the SDE Network service model.
    - The execution of OMOP and data linkage is an opportunity area for industry to support the SDE networks. Currently, adapting data to OMOP continues to be a somewhat manual process, inhibiting true interoperability.
- **Create sandpit environments with synthetic or anonymised data** not subject to GDPR that are accessible to industry and academic users as well as NHS research teams. Although the NHS is already trialling this across the Network, industry would benefit from early visibility and the involvement in the design process.
- While the SDE Network has already worked with Genomics England to map the NDRS onto a common data model, it can further **leverage the experience developed within the digital health, technology, and life science industry to map multi-modal data to open Clinical Data Model standards** (like OMOP and SDTM), as well as the experience in industry to leverage Generative AI to assist with this process.

Challenge 2: Federated approach

Given that the Data for R&D Programme's NHS Research SDE Network aims to achieve national coverage through a federation of sub-national Secure Data Environments, there are several challenges that must be addressed to ensure this approach is a success. SDEs each applying their own data models, or even different versions of common data models, will make aggregation of data for projects challenging. Crucially, without the ability to undertake national analysis federation is **likely to make international research more challenging**.

The SDE Network is looking to build upon the deliverables and lessons learnt from the DARE UK Phase 1 Driver Projects, as well as from industry partners to identify and implement the technologies needed to support federated analysis across the NHS Research SDE Network.

The intention is to use OMOP as a common data model where relevant for federated activities, starting with federated cohort discovery. Where OMOP is not appropriate, the Network is planning to identify use cases where other data models may be more relevant and test them in driver projects. They are also working with HDR UK to support a common researcher identity that can be trusted across federated data partners and their SDEs to enable a research project across multiple SDEs, complying with the approved datasets for the project and IG policies.

techUK's LSWG is highly supportive of these initiatives, but as it stands, a substantial increase in resources and funding will be necessary to deliver these at pace and achieve the transition from data sharing to access. The SDE Network should work with the technology industry to:

- Ensure that SDEs have a technically feasible interoperability approach and **investigate how best to achieve federated analytics** to provide national coverage and comparability of outputs at regional, national, and international levels. Global companies need to have rapid visibility of research viability in England (and the UK), and feasibility needs to be closely linked to delivery to ensure that outcomes are achieved and repeat business is secured.
- **Reach consensus about how, with what resources, to what data sets and in what priority order they are going to implement OMOP.**
- Planned to be delivered through the Health Data Gateway, **create and publish an open source SDE metadata** catalogue to improve data discoverability, enhance data understanding, encourage data exploration and analysis, and make it easier to collaborate and share knowledge.
- Continue to **support technical interoperability with equivalent standardised governance and commercial processes** through the IG and Ethics Working Group.
- Provide **clarity on expectations for projects which span boundaries** or cut broader than health care, as is being explored through an upcoming public deliberation.
- **Ensure all localities benefit equally** so as not to exacerbate health inequalities e.g. if certain SDEs cannot identify cohorts because they have not accessed GP data, will those patient groups not benefit from that research?
- Through a Bring Your Own Data and researcher accreditation approach (or other mechanism), **ensure common access and accreditation schemes** to enable collaboration between NHS with life science or technology industry partners who may be protective of their data assets

Challenge 3: Funding and Commercial Strategy

With an ambitious set of objectives, length of time allowed for delivery, total allocated funding, and uncertainty of budget allocation from year to year, SDEs may struggle to achieve their stated aims. Given that some SDEs are yet to enable live service, and most allocated funding supports the development of teams to deliver the service, insufficient funding may be available for technology, storage, and compute requirements. The running costs of SDEs to store and process data for over 5 million patients will be significant and could exceed programme funding. In-addition, SDEs need to be financially self-sustaining within a short timeframe but the commercial model to achieve this is not yet well understood.

The SDE Network should work with industry to:

- Consider novel ways of **analysing data at source** rather than duplicating and extracting bulk data into local data lakes or warehouses.
- Start small, focussing on data within areas that are most attractive to researchers and innovators (such as genomics) to generate additional revenue streams (through grants and industry funding), using this income to fund the next set of focus areas, eventually **becoming self-sustaining well in advance of any central funding coming to an end**.
- Utilise and **collaborate with existing academic and industry SDE providers** to create commercially sustainable models (e.g. fee-based model) which will increase revenue and potentially decrease cost for SDEs.
- Work collaboratively with industry to **establish other models beyond data access**, such as AI and predictive model development and establish consistent approaches to implement those in clinical workflow.
- **Map common insights coming from related data infrastructure projects in which industry are involved**, such as the Federated Data Platform and NHS App.
- Provide clarity regarding external vendors who currently provide, or plan to provide, competing services with NHS data; the commercial basis on which they do so; and create an equitable, transparent, and accessible mechanism for other vendors.

Challenge 4: Translating research to operational practice

In Spring 2023, Roland Sinker was asked by NHS England Chief Executive Amanda Pritchard to lead a review into how the NHS can partner with the life sciences and innovation sectors effectively. This work will seek to cover NHS support to enable and embed clinical research, and how the NHS can most effectively drive the adoption and spread of proven and cost-effective innovations to all patients who would benefit from them. There is a key opportunity for the SDE Network, system partners, and Health Innovation Networks to work with the NHS to get research learnings and outputs back into clinical practice.

The SDE Network should work with industry to:

- Seek advice on how to **scale research outputs in the health and care market**, with potential to co-invest and taking advantage of established processes and technology or develop new ways of implementing research outputs to deliver outcomes.
- **Identify technological mechanisms to translate research outputs into clinical practice**, for example cohort reidentification in digital clinical workflow tools, or the scaling of new predictive or prognostic models through the use of ML ops platforms.
- Acknowledge that **the SDE Network will only be part of the solution** and collaborate with industry to implement consistent approaches to returning actionable/executable knowledge back to front line services.

---

techUK welcome the continued industry engagement efforts of the Data for Research & Development Programme. We would like to offer industry's commitment to Secure Data Environments and willingness to engage at both a national and local level to ensure the UK reaps the widest possible benefits from SDEs. techUK is happy to facilitate specific roundtable or workshops to investigate specific suggestions in further detail.