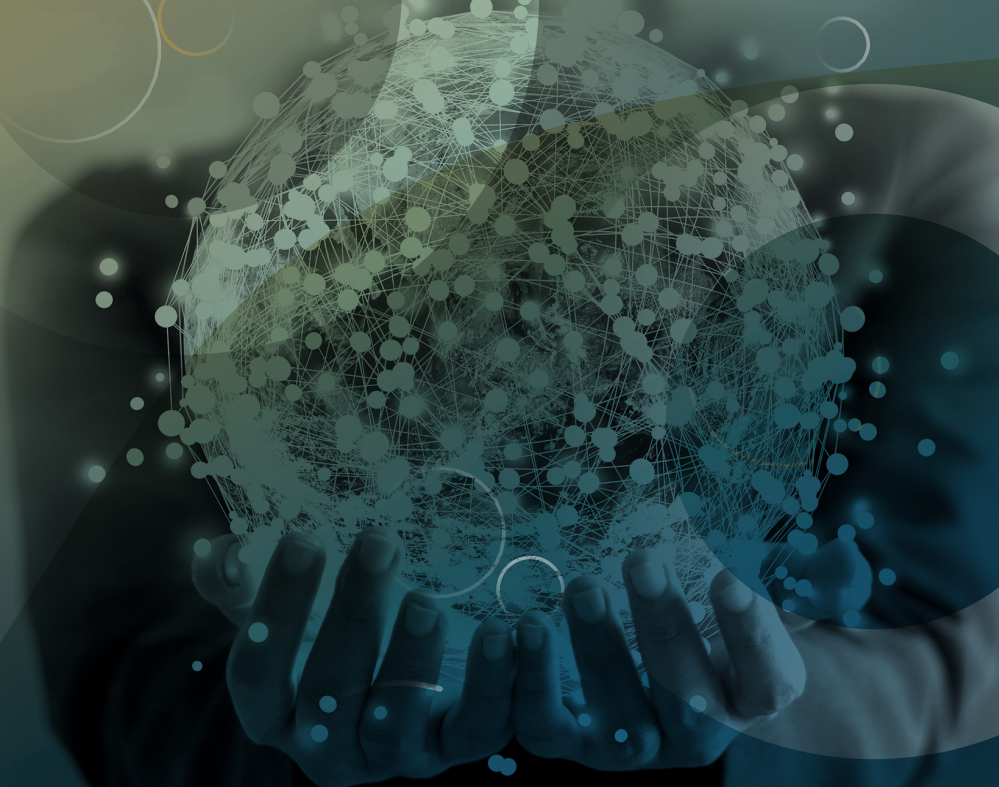


# Open Life Data Framework

Towards a framework for unlocking data flow and trusted data sharing for private sector derived health-relevant data.

How can we make our data work harder in a trustworthy way to help us live healthier for longer while reducing health inequalities



# CONTENTS

<b>Foreword</b>	3	4.2. How can we understand the data value chain and consent?	28
Who should read this report?	2	4.2.1. What is informed consent and consent management?	30
<b>1. Executive Summary</b>	4	4.2.2. What is data governance?	30
1.1. Vision, mission and scope	4	4.2.3. Can we help create a safe, secure open market?	32
1.2. A visual overview of data sharing processes	5	4.2.4. Can a sandbox approach aid learning and development?	33
1.3. Scope: what is this report's focus?	7	4.2.5. Can we triage rapid innovation while addressing risks and controls?	34
<b>2. Recommendations</b>	8	4.2.6. What role can Open Standards play in addressing consent-based data sharing?	36
<b>3. What questions do we aim to address?</b>	10	4.3. How can we link with national priorities, policies and related programmes?	37
3.1. Why is there an imperative to act on health-relevant data?	10	<b>5. Use cases - signposting the future</b>	42
3.2. How can technology and data support impact?	13	<b>6. Conclusions and Next Steps</b>	44
3.3. Why is 'healthy longevity for all' important?	15	<b>7. Acknowledgements</b>	45
3.4. Methodology: how were we guided?	16	<b>8. Appendix</b>	48
3.5. How can we (re)build trust and public confidence?	21	8.1. What are the potential benefits and opportunities?	48
3.6. The role of data: how can we assess its potential value?	22	8.2. What are potential risks, blockers and constraints?	49
3.7. How can we address commercial realities and evolving models?	24	8.3. What is data? The Data Spectrum	50
3.8. Are there potential regulatory levers?	25	8.4. Data maturity models	51
<b>4. Unlocking value</b>	26	8.5. Licensing frameworks, certification and assurance	52
4.1. What are the priority considerations for value creation?	26		
4.1.1. For people (citizens, patients)	26		
4.1.2. For the public sector	26		
4.1.3. For the private sector	27		
4.1.4. For the third sector	27		

## Who should read this report?

This report is intended for anyone interested in informing and shaping a publicly acceptable and trustworthy system to share data in our lives to improve our health and reduce health inequalities. This includes policy-makers in all areas affecting health (whether NHS, social care, transport, environment, education, work and pensions etc), clinicians and all health and care professionals, data scientists, leaders in business, charities, social enterprises and think tanks, social entrepreneurs, tech pioneers, ethicists, lawyers and general public.



# FOREWORD

The UK, along with the rest of the world, recognises that science and innovation are central to tackling the largest challenges we face as a planet. From climate change to ageing societies, and health emergencies like the pandemic, we must respond through innovation. Moreover, our economic recovery from the exceptional challenges we have recently faced must build on the UK's innovative foundations.

Our goal is to build on those foundations to grow a robust and agile economy, capable of the scientific & technological advances critical to building a greener, healthier and more prosperous future for the UK.

That is why we published “Build Back Better: our plan for growth” which focused on infrastructure, skills and innovation as the foundation for our economic recovery and growth. It is why, in July, we published the UK's Innovation Strategy, setting an ambitious programme of work across the innovation landscape to advance our vision of the UK as a global hub for innovation. And why, last month, the Government committed to increasing UK investment in R&D to £20bn by 2024/25, an increase of around a quarter over this Spending Review period and to take expenditure to £22bn by 2026/27.

The next decade will be an even greater time of change, challenge and opportunity. But, one thing is clear, data innovation will continue to play an increasingly important role in all our lives.

The Government is committed to unlocking the power of data. Our priority is to ensure an open, welcoming and secure environment where data improves life for people across the UK. Through the National Data Strategy, we are taking concrete steps to unlock the power of data and free up businesses and organisations to keep using data to innovate, experiment and drive growth.

The first mission of that strategy is to unlock the value of data across the economy, setting out how we will complement our high data protection standards by creating an environment where data is appropriately usable, accessible, and available. As part of this we are taking steps to secure a pro-growth and trusted data regime. To this end, we have been consulting on reforms to create a new, ambitious, and innovation-friendly data protection regime.

Over the summer, the Government has also engaged the health and care sector, and the public, on its draft data strategy for health and social care, Data Saves Lives. It sets out our vision for the future of health and adult social care data, recognising the need to realise the potential of data to drive transformation while maintaining the highest standards of privacy, ethics, and accountability.

Against this backdrop of work across Government, I warmly welcome this report produced by the APPG for Longevity. It makes a strong case for the role that sharing health-relevant data held within the private sector could play in providing powerful tools for individuals and communities to tackle numerous health challenges head-on, and to enhance and level-up health.

The examples it illustrates and its detailed contribution to this area are all the more welcome, having drawn on involvement from a wide variety of experts from across industry, academia, and the health sector itself. These findings provide valuable insights that will support the Government's work across a range of initiatives. Initiatives such as Smart Data, where the Government has already committed to legislate to enable schemes offering, with an individual's permission, the ability for individuals to share data held by private companies with trusted third parties to help them make sense and use that data in innovative ways.

Whilst this is an exciting area of innovation and technological drive, it is also a complex one which underscores the need for public trust and confidence. So, I want to thank the authors of this report, and the ongoing work of the APPG for Longevity, and encourage more voices to join our mission as we look to unlock the potential of data and realise this positive new future.

George Freeman MP  
*Minister for Science, Research and Innovation*



# 1. EXECUTIVE SUMMARY

## 1.1. Vision, mission and scope

**Our vision:** To enable consent-based, trust-enhancing and secure sharing of personal ‘health-relevant data’<sup>1</sup> from the private sector to improve and level up health and wellbeing.

**Our mission:** To create the enabling conditions for public and private sector actors to design and implement a framework that facilitates trustworthy data sharing for public benefit while unlocking innovation to increase healthy life expectancy by 5 years while minimising health inequalities.

The purpose of this report is both to stimulate conversation and encourage collaboration between public and private sectors. It seeks to explore the potential implementation of frameworks that can help provide access to private-sector derived, health-relevant data, and enable easier sharing with the public sector. Digitalised data is rapidly increasing in scope, source and connectivity. The core challenge that we aim to stimulate debate around is how the expansive and diverse explosion of data flows may be best managed to the benefit of the individual and the public, and supported by the public, private and third sectors.

Throughout, there is an emphasis on the need for innovation, continuous development and learning balanced with the need to address long-term stability. That focus must be placed to prioritise the delivery of measurable outcomes and to communicate clearly along the journey.

The scope of this work includes data that is generated by the private sector (both commercial and non-profit actors), such as that from apps and digital services.

It does not aim to address public sector data, for example, in electronic health records. To illustrate, we build on classifications in the paper “Privacy protections to encourage the use of health-relevant digital data in a learning health system”<sup>1</sup> and illustrate below.

### Definition of health-relevant data

**Category 1** - Health Care System Generated. Electronic medical record data, prescriptions, laboratory data—including molecular “omics” data, pathology images, radiography, payor claims data.

**Category 2** - Consumer Health and Wellness Industry Generated. Wearable fitness tracking devices, medical wearables (insulin pumps, pacemakers, medical or health monitoring apps, patient-reported outcome surveys, direct-to-consumer tests (eg DNA analysis) and treatments.

**Category 3** - Digital Exhaust Generated as a Byproduct of Consumers’ Daily Activities. Social media posts, Internet search histories, location and proximity data.

**Category 4** - Non Health Demographic, Social, and Economic Sources. Race, gender, income, credit history, employment status, education, level, residential zip code, housing status, census records, bankruptcy and other financial records, grocery store purchases, fitness club memberships, voter registration.

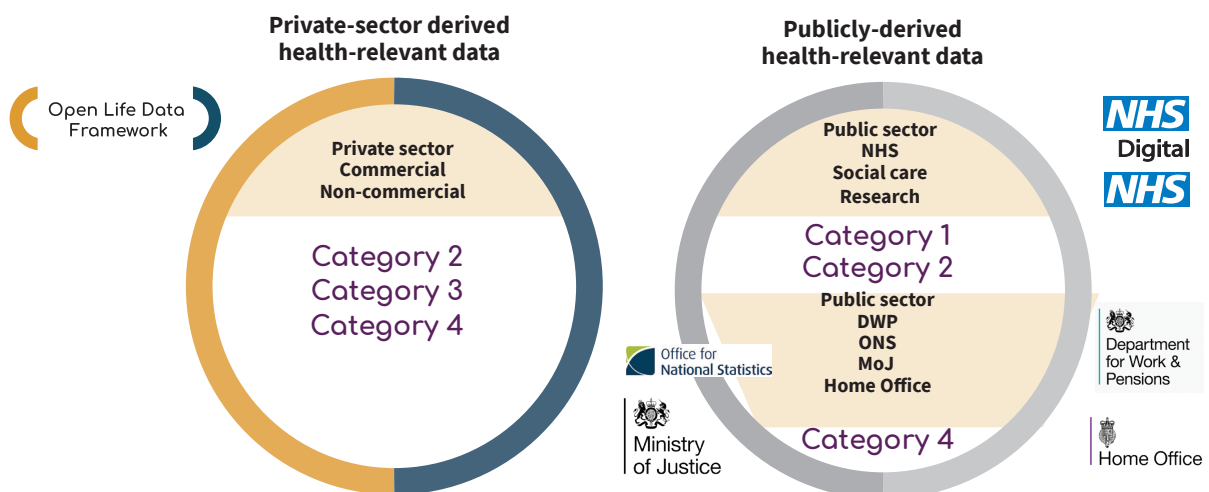


Figure adapted from reference 1 below.  
1. McGraw, D., Mandl, K.D. Privacy protections to encourage use of health-relevant digital data in a learning health system. *npj Digit. Med.* 4, 2 (2021).  
<https://doi.org/10.1038/s41746-020-00362-8>



For the avoidance of doubt, the framework discusses the principles of how health-relevant data may be **shared with consent** for a specific purpose to benefit individuals (e.g “can this company use this data for that purpose”) rather than exploring the specific purposes themselves, or how the processing for that specific purpose works inside the application (e.g “the AI inside the app works like this”). This is an important practical and legal point: all processing of personal data requires a legal basis, including providing **consent** for processing for a **specific** purpose.

It is also important to note that the processing of data is often the moment where Category 3 & 4 data become ‘health related’. For example, social media data is ‘just’ social media data until it’s processed to assess mental health status. At that moment of processing, it becomes health-related, and that’s where the particular health-related ethical and legal questions are opened for that use of data.

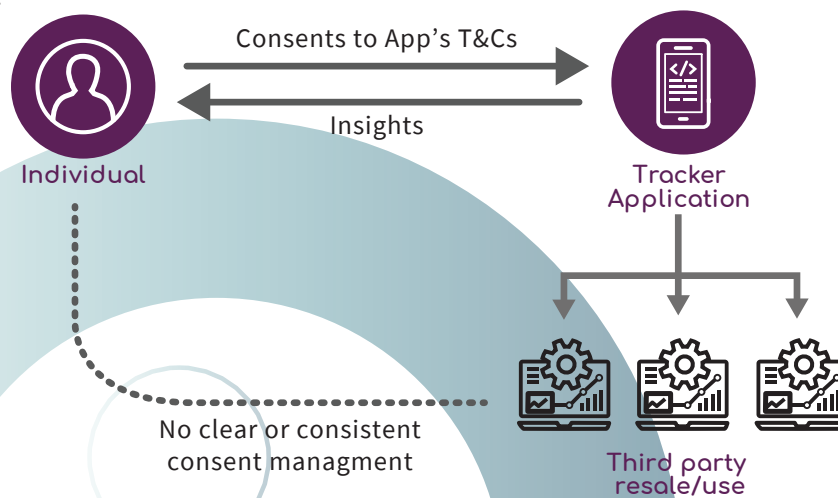
Further, the framework does not prioritise any specific data, use case, data codes (ontology, taxonomy or schema), nor seek to define any new ‘data standard’. The development of a framework may, in due course and through multi-stakeholder collaboration, help with the prioritisation of those areas, for which there are many examples to draw from.

The focus herein is about:

- A) how individuals and participants in the system can be **informed** about why, what and by whom data are being processed
- B) how they may **affect control** over the flow of data: how it can be turned on and turned off (noting that consideration needs to be given to individuals/ consumers who may lack time and social capital to understand and exercise all of their choices)
- C) what processes (operational, legal, technical) can be created to enable **cohesion and interoperability** across digitalised systems.

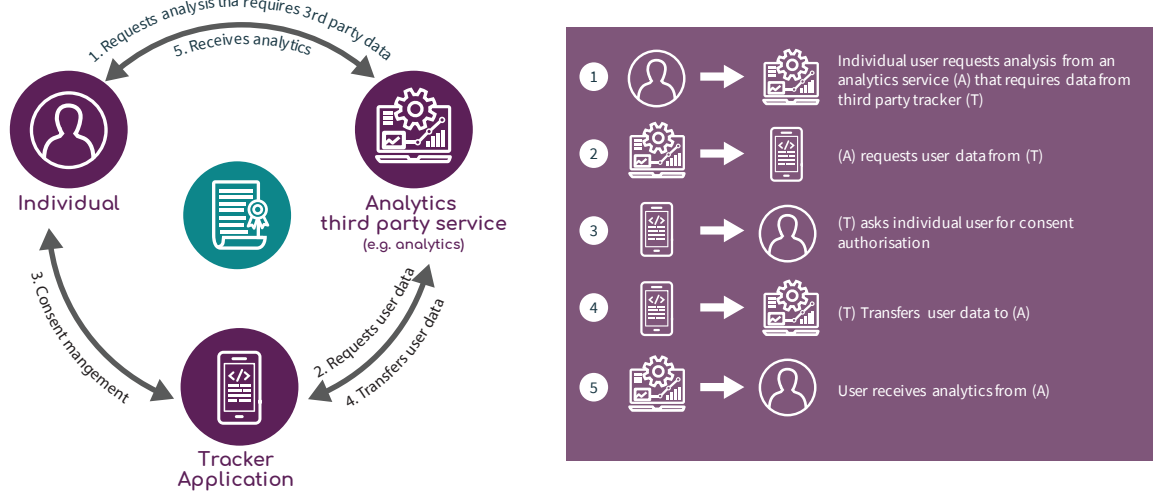
## 1.2. A visual overview of data sharing processes

**The situation today:** For a single application a user downloads and uses a tracker app. Data is stored and used by the app in a 1:1 contract between the user and the app company. It may include rights data resale/other use. Ongoing use is usually unclear and has no consistent consent mechanism.



In a framework that supports consent management, this relationship can be made both clear and consistent to the individual, and safer for the commercial vendors in the value chain, including making clear any potential onward sharing of data.

An example of consent -based data sharing between a tracker and analysis service

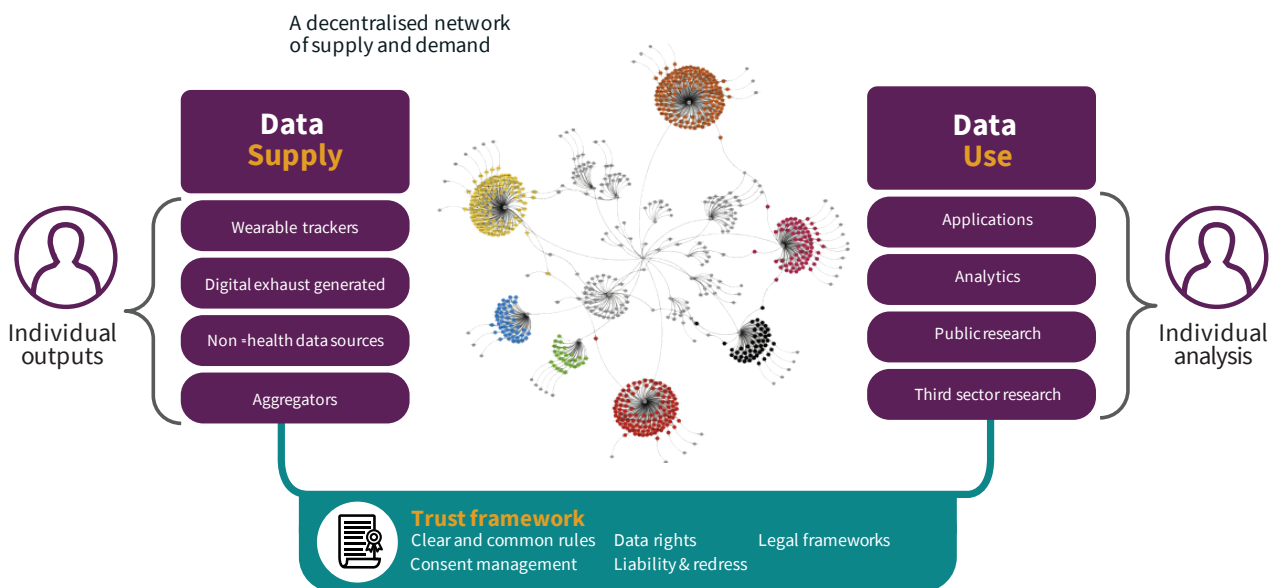


To achieve this at market-scale requires the development of a Trust Framework that can enable the definition of clear and common rules. Such rules make clear areas including, but not limited to:

- Data rights (e.g. GDPR)
- Legal frameworks (e.g. links to data processing agreements)
- Scope and assignation of liability
- Mechanisms for redress
- Consent and consent management
- User experience guidelines

In a decentralised and distributed market of data supply and usage, a common Trust Framework could help the development of a fair, secure and proportionate open market for data sharing.

### What is a Data Ecosystem?





### 1.3 Scope: what is this report's focus?

What the proposed Framework aims to cover, and what it does not address.

Aims to address	Does not address
<ul style="list-style-type: none"><li>● Private sector derived health-relevant data. Specifically, to address 85% of the determinants of our health that are not adequately addressed in current policies</li><li>● Applying open-source principles and move to open innovation and greater collaboration across all sectors and stakeholders with a role to play in enhancing and levelling up health</li><li>● Exploring how to unlock data flows between and across ALL health-relevant data (within and outside the healthcare system)- to derive maximum insights, unlock innovation, identify hidden populations, minimise data biases, focus on key health challenges and desired outcomes</li><li>● Aimed at making it easier for innovation to ignite, scale, and commercialise, harnessing data-intensive technologies and AI responsibly</li><li>● Maximising public benefit and enabling the private sector to flourish - within a trustworthy system to benefit individuals, and where commercial interests are proportionately balanced with benefits to the health and social care system and the wider public.</li><li>● Applying what has worked in other sectors to enable access and control of data while respecting privacy in line with evolving data protection regulation</li><li>● Guided by principles to ensure equitable and fair access to 'healthy lives for all' - recognising that value judgements, decisions and omissions will all have to be made to derive inferences about people's health from data</li><li>● Intended to inform, complement and support data initiatives underway by joining up the dots and identifying the gaps for maximum purpose-led innovation.</li></ul>	<ul style="list-style-type: none"><li>● Public sector derived health-relevant data (NHS and care data), reflecting 15% of the determinants of our health. While outside the scope of this report, the Framework nevertheless needs to be aligned with the principles and policies governing data use in the NHS and social care system- as good alignment between the two is critical to success.</li><li>● Setting policies for health-relevant data in the NHS and care system, which is covered by existing policies and guided by the NHS constitution, Common Law Duty of Confidentiality, Data Protection Act, and Right to Privacy, Health Research Authority's Confidentiality Advisory Group and Caldicott Principles.</li><li>● Providing recommendations on data linkage- this is outside of scope and would involve the complexities of genomics data</li><li>● Promoting 'privatisation of the NHS'. The aim is instead to promote wider, more transparent and controlled collaboration between private, public and third sectors rewarded by achieving the shared goal to maximise healthy lives equitably in line with UN Sustainability Development Goals (SDGs) and bringing 'Health' into Environmental, Social, and Governance (ESG) mandates</li><li>● Advocating openly licensing or openly publishing any individual's personal and private health information. Personal data is not and should not be made available as Open Data.</li></ul>

## 2. RECOMMENDATIONS

Following extensive research and consultation (see Section 3.4 page 16 for Methodology) , our key recommendations are as follows:

1. **Earn public trust through trustworthy practices via the development of a Trust Framework, a key enabler of the Open Life Data Framework.** It is important for data processors to be trustworthy to earn meaningful public trust. An independent body should oversee a transparent, multi-stakeholder process to develop the framework, with a mandate for open communication.
2. **Define clear and proportionate rules that can both enable and protect.** Rules must cover (at least) data rights, consent, transparency, liability and redress as well as operational mechanics to reduce transactional friction (e.g. open technologies).
3. **Call for legal obligations on Health Data Controllers (from National Government downwards) to engage in thorough public engagement processes where data is being shared.** This can build on methodology for public engagement now well established through, for example, Understanding Patient Data and OneLondon, while learning the lessons from Care.Data and the recent GP Data for Planning and Research (GPDPR) experiences.
4. **Enabled by the above, engage the public and health and social care community in a two-way dialogue on why and how data sharing can enhance and level up health, as part of a re-commitment to the government manifesto of 5 extra years of healthy life expectancy while minimising health inequalities.** A combination of public dialogue and literacy campaigns on a national scale could build public confidence. Public dialogue, on a large-scale or grassroots basis, would allow the public to engage with the topic from their perspective, for example, in community conversations to understand what people's hopes, fears, questions and concerns are, rather than being 'told' in a top-down way. On the other hand, a national health, data and information literacy campaign on why data matters and how data saves lives and promotes health and care, backed up by real examples that evidence public benefit, could inform the public debate.
5. **Assign an independent body to be accountable for health-relevant data standards and frameworks.** The assignment of responsibility to NHS/NHSX and social care and an independent body (the health equivalent of FCA) to both co-lead and have joint accountability for developing, implementing and monitoring an Open Life standard. This independent body could be modelled on the Life Sciences Council to coordinate activity across different bodies (including NHSX, D, E, I, HDRUK, ONS, DCMS, Data Guardian etc) to align standards, frameworks, approaches and ensure accountability. A clear governance structure and statement of public interest will help to unblock barriers.
6. **Set up a new role of National Data Coordinator to assist in standardised integration of health-relevant data in the context of Integrated Care Systems (ICs) and Trusted Research Environments (TREs) accountable to an independent body (as above).** This role should have a horizon scanning and future looking function, working with relevant NHS stakeholders, including the move to networked TREs and federated analytics.
7. **Ensure the switch towards a national grid of TREs is funded with the investment required to build and operate a new system,** recognising there will be significant transitional and training costs involved.
8. **Fund use cases and sandbox projects to demonstrate the principles of the Open Life Data Framework.** Create a diverse suite of exemplars that demonstrate best practice in trustworthy and responsible data processing, while delivering public benefit, value for money and proportionate risk/reward for all stakeholders (refer to Section 4 for more detail) in the creation of an equitable healthy longevity industry.
9. **Develop and refine commercial models and risk/reward scenarios** using real-world use cases (including from both NHS and social care), in line with the NHSX draft Health and Social Care Data Strategy and Centre for Improving Data Collaboration (CDIC) Value Sharing Framework



and Guidance, guided by ethical frameworks being piloted by the Centre for Data Ethics and Innovation, Ada Lovelace, and other groups like BSI, and aligned to ONS Health Index, and ESG mandates to enhance and level up health.

10. **Support and reinforce ongoing data-led initiatives and reforms to level up health.** For example, support data-led initiatives of National Food Strategy recommendations, Levelling Up Taskforce and the Office for Health Improvement and Disparities.
11. **Ensure alignment with ongoing policy developments on data, Artificial Intelligence, international trade and agile regulatory reform.** Ensure there is coordination between developments underway with National Data Strategy, National AI Strategy, draft Health and Social Care Data Strategy, NHSX AI Strategy, and Taskforce on Innovation, Growth and Regulatory Reform (TIGRR) to ensure joined-up approach.

Note the above recommendations can apply to England and the devolved nations as appropriate. For more detail on how the Framework will benefit the general public, and private, public and third sectors, please read Section 4.1 (pages 26-27). For next steps, refer to Section 6 (page 44).

*Freeing up health and care data currently locked in commercial 'silos' in ways that mean usefully structured data can be shared with those providing people with actual health and care services would be a good thing. You might start with the private and various local authority social care providers, and the private hospitals, and then maybe Google/FitBit, Apple Health Kit, Amazon Web Services Health, Microsoft HealthVault, etc. And once the Open Life Data Framework has demonstrated it can provide real value in and along care pathways – where, for example, the consent problem set should be far more tractable – making such data available for secondary uses may then become more of a viable option.*

Independent contributor

*We did an exercise to pull out data from companies in the UK data economy on Environmental, Social and Governance (ESG) criteria and their contribution to the Sustainable Development Goals (SDG) and what we found was that there is a big gap to confidently estimate how well companies are working towards the SDGs - you need standards.*

Lisa Allen, Head of Consultancy for Data Programmes, Open Data Institute

*We need the UK equivalent of the US Office of the National Coordinator of Health Information Technology to drive cross-Government cohesion at pace. The UK data-sharing and health system contexts should be more innovation-friendly than in the US where provider integration for population health is difficult. Yet the UK components sit across different Government departments including DHSC, MHCLG, DfE and DCMS. A world leading national grid of intelligence-led integrated care systems is entirely possible, but only if there is a cross government department mandate and slick operation.'*

Professor Iain Buchan, Chair of Public Health and Clinical Informatics and Executive Dean of Institute of Population Health, University of Liverpool

*The reservoir of trust is already with the NHS- if we are dealing with health-relevant data the overseer and arbiter role for accountability needs to be a health-focussed organisation independent from central government.'*

Aidan Peppin, Senior Researcher, Ada Lovelace Institute

## 3. WHAT QUESTIONS DO WE AIM TO ADDRESS?

### 3.1. Why is there an imperative to act on health-relevant data?

The pandemic has highlighted the poor state of the UK's health compounded by significant health inequalities that will continue to worsen unless wider systemic issues are addressed.<sup>2</sup> There is a growing appreciation that levelling up the economy is not possible without levelling up health.<sup>3</sup>

It must be acknowledged that health data is also politicised. To 'build back better'<sup>4</sup> health and care needs to be seen as an asset to invest in, with long term returns measured by closing the gap in health inequalities, reducing demand on the health and care system, and growing health and wellbeing of populations with related increases in labour productivity and, ultimately, economic growth.

Measuring what matters gets done, yet current performance metrics in the healthcare system are not geared enough to incentivise prevention, and we are seeing a healthcare system becoming increasingly unsustainable post-pandemic. We need to re-gear incentives as part of wider systemic change to link health to wealth.

While government health priorities, policies, and funding continue to focus on the NHS and care system to bail out the nation's health in post-Covid recovery, there is hope that the new Office for Health Improvement and Disparities (OHID) can implement a more health-in-all- policies approach to address the 85% of the factors determining our health that lie outside the formal health and care system.<sup>5</sup>

Initiatives like the wearables pilot announced in October to nudge and reward people to move more (initiated by OHID) may help tackle obesity-related illnesses that cost the NHS £6 billion a year, and are exacerbated by COVID-19 disproportionately affecting people who are overweight.<sup>6</sup> Other innovative schemes are already underway like councils rewarding citizens with 'civic dollar' rewards for spending time in parks<sup>7</sup> recognising that green spaces lead to better health.<sup>8</sup>

The pandemic has shown how valuable data is to protect our health.<sup>9</sup> The public is increasingly aware of how data can improve health at individual and population health levels- from the lightning-speed innovation and deployment of vaccines, to the virtual GP consultations that appeared literally overnight to the therapies such as dexamethasone that were discovered through research.

We need to maintain and indeed accelerate this pace and sense of urgency to improve and level up our health. But to do this we must focus more on the 85% of the determinants of our health and we need to seek solutions outside the formal healthcare system.<sup>10 11</sup> The development of the ONS Health Index will help focus public debate and policy attention across government on a broad concept of health as an asset and a key measure of success.<sup>12</sup>

Other indices, like the Urban Health Index, show how layering data relating to different social and environmental indicators can build a better picture of the circumstances of people living in urban

2. Green, D, Filkin G, Woods T. Our unhealthy nation. *The Lancet Healthy Longevity*, Volume 2, Issue 1, 2021, Pages e8-e9, [https://www.thelancet.com/journals/lanhl/article/PIIS2666-7568\(20\)30062-3/fulltext](https://www.thelancet.com/journals/lanhl/article/PIIS2666-7568(20)30062-3/fulltext)

3. *Levelling Up Health. Report by APPG for Longevity. April 2021.* <https://appg-longevity.org/levelling-up-health>

4. *Build Back Better: Our Plan for Health and Social Care, 15 September 2021,*

<https://www.gov.uk/government/publications/build-back-better-our-plan-for-health-and-social-care/build-back-better-our-plan-for-health-and-social-care>

5. Buck D, Baylis A, Dougall D, Robertson R, *The Kings Fund. A vision for population health: towards a healthier future.* The Kings Fund, 2018. <https://www.kingsfund.org.uk/sites/default/files/2018-11/A%20vision%20for%20population%20health%20online%20version.pdf>

6. *New pilot to help people eat better and exercise more, 22 October, 2021* <https://www.gov.uk/government/news/new-pilot-to-help-people-eat-better-and-exercise-more>

7. <https://www.telegraph.co.uk/technology/2021/04/03/council-bribes-locals-civic-dollars-spend-time-parks>

8. Ivers L. *Green Spaces Are a Necessity, Not an Amenity. How Can Cities Make Them Accessible to Everyone?* 11 April 2021. <https://www.discovermagazine.com/health/green-spaces-are-a-necessity-not-an-amenity-how-can-cities-make-them>

9. Stevenon A et al. *Strengthening social care analytics during the pandemic and beyond*, Health Foundation, 5 May 2020. <https://www.health.org.uk/news-and-comment/blogs/strengthening-social-care-analytics-during-the-pandemic-and-beyond>

10. Green D, O'Shaughnessy J, Starks G, Sloggett R, Buchan I, Woods T. *Open Life Data to support healthy longevity for all.* *Lancet.* 2021: [http://doi.org/10.1016/S2666-7568\(21\)000081-7](http://doi.org/10.1016/S2666-7568(21)000081-7)

11. *The Data will See You Now.* Ada Lovelace Institute, October 2020. <https://www.adalovelaceinstitute.org/report/the-data-will-see-you-now/>

12. *The Health Foundation's response to Office for National Statistics consultation on the 'beta' version of the Health Index,* [https://www.health.org.uk/sites/default/files/2021-04/final\\_ons\\_health\\_index\\_consultation\\_hf\\_response.pdf](https://www.health.org.uk/sites/default/files/2021-04/final_ons_health_index_consultation_hf_response.pdf)



areas and how their environment impacts their health.<sup>13</sup> Global initiatives such as the Partnership for Healthy Cities<sup>14</sup> are already underway like Vancouver's Healthy Cities data dashboard, which launched in April 2021, tracking progress against 23 health and wellbeing indicators and 12 shared targets.<sup>15</sup> These are all examples that link ecological and social outcomes recognising humans live in a wider ecosystem.<sup>16</sup>

These innovative projects show how we need to demonstrate the value of transparent, consent-based, secure sharing data with the public. More and more people are getting used to having the NHS app on their phones<sup>17</sup> to access their medical records and demonstrate their vaccination status -but the pandemic has also shown the tensions on using data to save lives and address inequalities while protecting personal privacy.<sup>18 19 20</sup>

The Open Life Data Framework proposed will focus on making it easier to harness, share and unlock innovation with 'health-relevant data'<sup>21</sup> (see Figure 1) in a trustworthy, publicly acceptable way to benefit individuals. Privacy concerns need to be addressed but so does the flow of data to improve population health and ensure disparities in health outcomes are tackled as part of a learning healthcare system.<sup>22</sup>

It is important when discussing health and social care data, that both health and social care providers or experts are involved in decisions. The right balance needs to be struck through an informed debate. It is hoped the Framework will close the gap in understanding health-relevant data and why sharing it matters so much to enhance and level up health and wealth, as part of a new social contract.<sup>23</sup> This is particularly timely as the pace quickens with health and care reforms to integrate health and care through greater collaboration, agile regulatory reform, and government ambitions to lead in science, technology and innovation underpinned by national strategies in data and AI<sup>24</sup> that promote the use of Trustworthy Research Environments (TREs).

*The NHSX is getting increasing calls for how to factor standardised integration of health-relevant data, particularly at an Integrated Care System (ICS) level, to address health inequalities and pandemic recovery.*

Lisa Murphy, Programme Manager, NHSX CDIC

*An emerging area which needs further consideration is the data gathered by personal apps, prescribed or not by the health and care system. Who controls access to this data remains obscure and often, in order to use an app, rights to know or control the use of data have to be waived. If the work proposed includes integrating this data with NHS or social care data, this needs to be clarified alongside what systems would be in place for managing the data which replicate the controls on NHS data...*

The National Data Guardian

13. Urban Health Index, <https://urbanhealth.org.uk/insights/data/urban-health-index-uhi-for-lambeth-and-southwark>

14. <https://partnershipforhealthycities.bloomberg.org/>

15. Healthy cities take centre stage in urban planning, <https://cities-today.com/healthy-cities-take-centre-stage-in-urban-planning/> 28 October 2021

16. Lydia P. Olander, Robert J. Johnston, Heather Tallis, James Kagan, Lynn A. Maguire, Stephen Polasky, Dean Urban, James Boyd, Lisa Wainger, Margaret Palmer, Benefit relevant indicators: Ecosystem services measures

17. What if smartphones became personal health assistants? The Economist, 3 July 2021, <https://www.economist.com/what-if/2021/07/03/what-if-smartphones-become-personal-health-assistants>

18. Data Sharing in a Pandemic: Three Citizens' Juries Juries' Report, National Institute for Health Research, [https://www.arc-gm.nihr.ac.uk/media/Resources/ARC/Digital%20Health/Citizen%20Juries/12621\\_NIHR\\_Juries\\_Report\\_ELECTRONIC.pdf](https://www.arc-gm.nihr.ac.uk/media/Resources/ARC/Digital%20Health/Citizen%20Juries/12621_NIHR_Juries_Report_ELECTRONIC.pdf)

19. Public attitudes to COVID-19, technology and inequality: a tracker. Ada Lovelace Institute, 5 March 2021. <https://www.adalovelaceinstitute.org/feature/public-attitudes-covid-19-technology-inequality-tracker/>

20. The Data Divide. Public attitudes to tackling social and health inequalities in the Covid-19 pandemic and beyond. Ada Lovelace Institute. 25 March 2021, [https://www.adalovelaceinstitute.org/wp-content/uploads/2021/03/The-data-divide\\_25March\\_final-1.pdf](https://www.adalovelaceinstitute.org/wp-content/uploads/2021/03/The-data-divide_25March_final-1.pdf)

21. McGraw, D., Mandl, K.D. Privacy protections to encourage use of health-relevant digital data in a learning health system. *npj Digit. Med.* 4, 2 (2021). <https://doi.org/10.1038/s41746-020-00362-8>

22. Ainsworth J, Buchan I. Combining Health Data Uses to Ignite Health System Learning. *Methods Inf Med.* 2015;54(6):479-87. doi: 10.3414/ME15-01-0064. <https://pubmed.ncbi.nlm.nih.gov/26395036/>

23. The Social Contract for Data. World Bank. 2021. <https://wdr2021.worldbank.org/stories/the-social-contract-for-data/>

24. UK AI Council Roadmap. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/949539/AI\\_Council\\_AI\\_Roadmap.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/949539/AI_Council_AI_Roadmap.pdf)

The framework we are proposing aims to improve the ability for people to share health-relevant data in a manner that is secure, fair and protected, addressing needs of the public voiced in numerous citizen research studies and public forums. <sup>25 26 27 28 29 30</sup>

*Over the last few years, and accelerated by Covid, the public debate has moved on from 'how do you feel about your data being used in these different ways' to 'how do you expect it to be accessed and what controls do you want to see in place?'*

Michelle, Mackie, Head of the Qualitative Research and Engagement Centre, Ipsos Mori

The framework aims to address the needs of the individual while making clear the process of consent and consent management. It must also make issues clear, including commercially enforceable processes of liability transfer, modes of redress and related issues that are blockers to the effective and fair use of data by the private, public and third sectors.

Without proactive engagement in creating a scalable, pragmatic framework, there is a risk the organisations will either gain greater access to personal data without proportionate and due care of the citizen (their customer) or that access to data will be overly restricted such that the benefits may not be realised for those in need.

The balance we seek to achieve is the implementation of controls such that those involved can assess and address what they determine to be appropriate risk or reward (via which tangible benefits can be realised).

While many people are happy to share their data for personal and wider societal benefit,<sup>31</sup> health data represents a highly contentious and emotive area, especially when government and commercial interests are involved. The primary purpose of this report is to explore by example, and propose potential models that are appropriate to address the realities of sharing personal data—given that this is something that is already happening and will continue to scale.

We propose that such processes and controls must be codified (i.e. that data governance and consent management themselves must be able to be expressed in code and data) in a scalable, repeatable and consistent manner.

We further propose that creating common guidance and appropriate rules that can be applied across the private sector in such a manner will ultimately reduce risks, increase transparency, improve citizen control and enable regulatory and policy instruments to be applied- in a publicly acceptable and trustworthy system.

Such common guidance and rules can potentially be adopted through voluntary codes of practice, codified through contractual agreements and/or regulated via appropriate bodies.

25. *The One-Way Mirror: Public attitudes to commercial access to health data*, March 2016

<https://www.ipsos.com/sites/default/files/publication/5200-03/sri-wellcome-trust-commercial-access-to-health-data.pdf>

26. *Public dialogue on the ethics of data science in government*, May 2016 <https://www.ipsos.com/sites/default/files/2017-05/data-science-ethics-in-government.pdf>

27. *Future data-driven technologies and the implications for use of patient data Dialogue with public, patients and healthcare professionals*, Report for the Academy of Medical Sciences, November 2018 <https://acmedsci.ac.uk/file-download/6616969>

28. *Public deliberation in the use of health and care data*, OneLondon Citizen Summit, 2020 March <https://www.onelondon.online/wp-content/uploads/2020/06/Public-deliberation-in-the-use-of-health-and-care-data.pdf>

29. *Discover NOW Citizen Advisory Group* <https://imperialcollegehealthpartners.com/portfolio/discover-now-citizen-advisory-group-cag/>

30. *A public dialogue on genomic medicine: time for a new social contract?* <https://www.ipsos.com/sites/default/files/ct/publication/documents/2019-04/public-dialogue-on-genomic-medicine-full-report.pdf>

31. *Quick guide to explaining how patient data is used*, *Understanding Patient Data*, October 2021, <https://understandingpatientdata.org.uk/sites/default/files/2021-10/Resource%20quick%20guide%20v%202.0.pdf>



## 3.2. How can technology and data support impact?

Our collective action challenge is to balance values while balancing value.

Data and technology do not themselves create change, they are resources and tools. Better access to data and technology can enable relevant solutions to be built that can support changes in health, care and wellbeing. There are new opportunities emerging: new business models and new products, and new ways to build public trust through distributed models.<sup>32</sup> There are new ways for organisations to engage with their customers and ways for the public and private sectors to collaborate to create public benefit.

### What is public benefit?

The National Data Guardian (NDG) advises and challenges the health and care system to help ensure that citizens' confidential information is safeguarded securely and used properly. The NDG recently published research to understand and define 'public benefit'<sup>33</sup> in the context of NHS health and care data. The findings show what matters most to the public to ensure that health and care data is used in ways that benefit people and society are:

- Equitable distribution of benefits of data use in health and social care with safeguards to protect against discrimination and geographic disparities
- Identifiable and sensitive data should be treated with the utmost care, if it is, it has the potential to bring public benefit. Data was perceived as being particularly sensitive if it is of a personal nature, such as genomics or mental health data, or because greater care is needed in its interpretation, such as qualitative data
- Safeguards and provisions in place to protect society from data manipulation, where the outputs from the data use could be interpreted in different ways, for example, to achieve political or financial ends. This includes publication of statements of data users' credentials and sources of funding
- Public benefit must outweigh profit with profitable uses of data rigorously scrutinised for demonstrations of public benefit before access is granted. There is a recognition that data use in this context can enable health and social care improvements and innovations
- Being ambitious for health and care data use - to realise public benefit from global collaboration; exploratory research driving breakthroughs; and using profit for new developments, such as drugs, treatments and services.

This builds on guidance from WHO on data as a public good, in which digital public goods as: 'open source software, open data, open AI models, open standards and open content that adhere to privacy and other applicable laws and best practices, do no harm, and help attain the Sustainable Development Goals.'<sup>34</sup>

The Covid-19 pandemic has accelerated some of these changes. Many organisations are now bringing forward their digital transformation plans to respond to new operational challenges and support people.

33. *Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data*, National Data Guardian, 14 April 2021, <https://www.gov.uk/government/publications/putting-good-into-practice-a-public-dialogue-on-making-public-benefit-assessments-when-using-health-and-care-data>

34. *Health Data as a Global Public Good*. Health Data Governance Summit. 30 June 2021. [https://cdn.who.int/media/docs/default-source/world-health-data-platform/events/health-data-governance-summit/preread-2-who-data-governance-summit\\_health-data-as-a-public-good.pdf?sfvrsn=2d1e3ad8\\_8](https://cdn.who.int/media/docs/default-source/world-health-data-platform/events/health-data-governance-summit/preread-2-who-data-governance-summit_health-data-as-a-public-good.pdf?sfvrsn=2d1e3ad8_8)

In parallel, the burdens on health and care are increasing. We need to find new ways to help deliver appropriate care at the point of need and also to prevent and detect diseases earlier to reduce demand. Digital technologies and data can help but they also introduce new challenges and inequities that must be addressed. Careful consideration of what is proportionate is essential and, ultimately, solutions must benefit their users, communities and society without compromising on privacy or generating new harms.

The scope of this work is not to determine what specific data are useful (although examples will be used to illustrate). Instead, its focus is on the framework for consent-based sharing. There are precedents to build upon. Open Banking<sup>35</sup>—in which the UK has led the world—is now used by 3 million people<sup>36</sup> across the country, giving customers access to their data to rebalance the market in their favour. By enabling data portability and creating greater competition, there has been greater innovation from new entrants and incumbents (over 300 fintechs and innovative providers have joined the ecosystem since the start of Open Banking), leading to greater convenience and engagement from customers regarding their finances.<sup>37</sup> Based on the success of that initiative the Government, regulator (FCA) and industry are exploring how to expand it across ‘Open Finance’.<sup>38</sup>

*Transparency and accountability are increasingly important for the safe, ethical and compliant use of data and machine learning. As algorithmic decision-making becomes more widespread, it needs to be understandable to build public confidence and ensure it is bias-free, well-governed and operating in consumers’ and markets’ interests.<sup>39</sup>*

#### FCA Business Plan

In parallel, the Government, regulator (Ofgem) and industry are exploring the development of ‘Open Energy’ and the potential to apply similar principles across that sector.<sup>40</sup> These initiatives are also connected in their architecture, design and operational principles to a cross-sector approach to Smart Data.<sup>41</sup>

This report explores the application of these principles and practices to commercially derived health-relevant data and its potential impacts on both ‘inside’ and ‘outside’ our primary health and care systems. It aims to aid the development of understanding:

1. where there may be benefits to copying existing approaches;
2. where there may be material gaps in user needs, expectations and applications;
3. whether alignment may bring substantial benefits for the whole value chain, through standardisation of approach, and the maximisation of cohesion and interoperability;
4. what commercial models can best stimulate innovation and encourage collaboration while maximising public benefit.

Specifically, it explores the potential benefits, opportunities, risks and threats related to such an approach. To help ‘bring the story to life’ use cases are used to better understand and illustrate potential outcomes.

35. What is Open Banking, <https://www.dgen.net/0/open-banking/>

36. Open Finance Feedback Statement, Financial Conduct Authority, March 2021 <https://www.fca.org.uk/publication/feedback/fs21-7.pdf>

37. Three years since PSD2 marked the start of Open Banking, the UK has built a world-leading ecosystem

<https://www.openbanking.org.uk/news/three-years-since-psd2-marked-the-start-of-open-banking-the-uk-has-built-a-world-leading-ecosystem/>

38. <https://www.fca.org.uk/publication/business-plans/business-plan-2019-20.pdf#chapter-4>

39. <https://www.fca.org.uk/publications/corporate-documents/our-business-plan-2019-20>

40. <http://openenergy.org.uk>

41. <https://www.gov.uk/government/publications/smart-data-working-group-spring-2021-report>

### 3.3. Why is 'healthy longevity for all' important?

The UK (and the world) is undergoing a demographic shift with an increasingly ageing population linked to an increased risk of poor health. Without new approaches, costs of age-related disease and social care will escalate, and the quality of life will worsen for a significant sector of the population.

As part of this, attention needs to be re-ignited on the current government manifesto commitment to increase healthy life expectancy by 5 years while minimising health inequalities by 2035. But this will only be achieved as a cross-sector pan-societal endeavour involving public, private and third sectors, and owned by all citizens.

Health can be improved across the life course through enhanced prevention and even in later life with simple drug therapies.<sup>42</sup> Interventions informed by our understanding of the biology of ageing and the exposome<sup>43</sup> are applicable to chronic health conditions arising from lifestyle and socioeconomic factors including obesity<sup>44</sup> and diabetes, and acute infectious diseases including COVID-19.<sup>45 46</sup> Adding just one extra year of healthy life expectancy (HLE+1) will have enormous financial benefits estimated at \$38 trillion in the USA.<sup>47</sup>

Enabling data sharing across the life course in multiple sectors will accelerate our understanding of drivers of age-associated poor health and identify preventative health strategies and therapies with the potential to improve healthy life expectancy and to level up health across the UK. There are exciting developments in areas like ageing biomarker research,<sup>48 49 50 51 52</sup> for example, and finding the cure for Alzheimer's Disease is one of the potential prizes.<sup>53</sup>

Through a recommitment to the government manifesto goal of 5 extra years of healthy life expectancy while minimising inequalities we propose embedding a healthy longevity focus in the National Data Strategy,<sup>54</sup> National AI Strategy and draft Health and Social Care Data Strategy to simplify data use by researchers, and developers of AI and other cutting-edge technologies and promote wider collaboration between the public and private sectors to enhance preventative health strategies linked to the government's Build Back Better recovery strategy.<sup>55</sup>

While recognising the significant issues of trust and concerns over health data flowing between systems, agile regulatory reforms<sup>56 57</sup> could facilitate a move towards an open innovation environment for frictionless sharing of health-relevant data within and outside the NHS and care system. This could maximise the opportunities for citizens and for society, as well commercial returns for UK plc by supporting innovation and the SME sector, spurring new trade deals, and creating a vibrant data-driven healthy longevity marketplace (including new financial investment and insurance products, cryptocurrency exchange, Data Trusts, citizen data donation etc).

If we get it right, 'healthy longevity' could emerge as a high-growth discipline with significant commercial and societal potential and presents an exciting opportunity for the UK to become a world leader.

42. Cox LS *The Health of the Nation – a strategy for healthier longer lives.* p72-80 <https://appg-longevity.org/health-of-the-nation>

43. Vineis P et al. *What is new in the exposome?* *Environment International* 2020;1;143:105887, ISSN 0160-4120, <https://www.sciencedirect.com/science/article/pii/S0160412020318420?via%3Dihub>

44. *Citizen science to combat obesity.* <https://www.infino.me/>

45. Cox, LS and Lord, JM (2021) *Coronavirus: Targeting aging cells improves survival* *SCIENCE* 373(6552);281-282 DOI: 10.1126/science.abi4474

46. 'Inflammation clock' can reveal body's biological age: Using machine learning, researchers created a tool that might help doctors improve people's healthy lifespan. *Nature*, 13 July 2021. <https://www.nature.com/articles/d41586-021-01915-x>

47. Scott, A.J., Ellison, M. & Sinclair, D.A. *The economic value of targeting aging.* *Nat Aging* 1, 616–623 (2021). <https://doi.org/10.1038/s43587-021-00080-0>

48. *Biomarkers of human mortality.* <http://mortalitypredictors.org/>

49. Zhavoronkov A, Bischof E, Lee K. *Artificial intelligence in longevity medicine.* *Nature Aging* 2021;1;5-7. <https://www.nature.com/articles/s43587-020-00020-4>

50. Peto MV, De la Guardia C, Winslow K, Ho A, Fortney K, Morgen E. *MortalityPredictors.org: a manually-curated database of published biomarkers of human all-cause mortality.* *Aging (Albany NY)*. 2017 Aug 31;9(8):1916-1925. doi: 10.18632/aging.101280. PMID: 28858850; PMCID: PMC5611985.

51. *Digital Ageing Atlas.* <https://ageing-map.org/>

52. *Science and Technology Select Committee 1st Report of Session 2019–21 HL Paper 183 Ageing: Science, Technology and Healthy Living*

53. *Alzheimer's Research UK partners with global leaders including Bill Gates to revolutionise detection of neurodegenerative diseases, Alzheimers Research UK,* <https://www.alzheimersresearchuk.org/alzheimers-research-uk-partners-with-global-leaders-including-bill-gates-to-revolutionise-detection-of-neurodegenerative-diseases/>

54. *UK National Data Strategy* <https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy#executive-summary>

55. *Build Back Better Plan for Growth,* <https://www.gov.uk/government/publications/build-back-better-our-plan-for-growth>

56. *UK unveils post-Brexit global data plans to boost growth, increase trade and improve healthcare*

26 August 2021 <https://www.gov.uk/government/news/uk-unveils-post-brexit-global-data-plans-to-boost-growth-increase-trade-and-improve-healthcare>

57. *UK Innovation Strategy Leading the future by creating it.* July 2021 [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1009577/uk-innovation-strategy.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1009577/uk-innovation-strategy.pdf)



### 3.4. Methodology: how were we guided?

The Open Life Data Framework was a recommendation in the APPG for Longevity's report, *The Health of the Nation*,<sup>58</sup> published in February 2020. Its rationale was published in the *Lancet*, *Open Life Data to Support Healthy Longevity for All*,<sup>59</sup> on the same day the APPG published *Levelling Up Health*<sup>60</sup> on 9 April 2021.

The APPG formed an expert working group comprising 45 individuals (see Acknowledgements). Two task meetings took place in November 2020 which informed the APPG for Longevity response submitted to the National Data Strategy consultation that closed on 9 December 2020. Three working group meetings took place between February-July 2021 to inform a wider consultation that ran between August-October 2021 based on an FAQ document. In all, 45 people participated in the working groups, 55 completed the survey and 20 took part in detailed one-to-one interviews.

Three use cases were peer-selected to illustrate potential issues to address (see boxed sections). The use cases put forward for selection highlighted that there is a very broad spectrum of data maturity and data literacy, and that many solutions are still highly 'technology-led' (i.e. driven by what might be possible rather than what user needs might be addressed through enabling technologies). Collectively, while many addressed large parts of the data value chain, they demonstrate a clear need for holistic training to enhance understanding of how to unlock the potential of data use, including legal, operational and technical areas.

58. *The Health of the Nation – a strategy for healthier longer lives* <https://appg-longevity.org/health-of-the-nation>

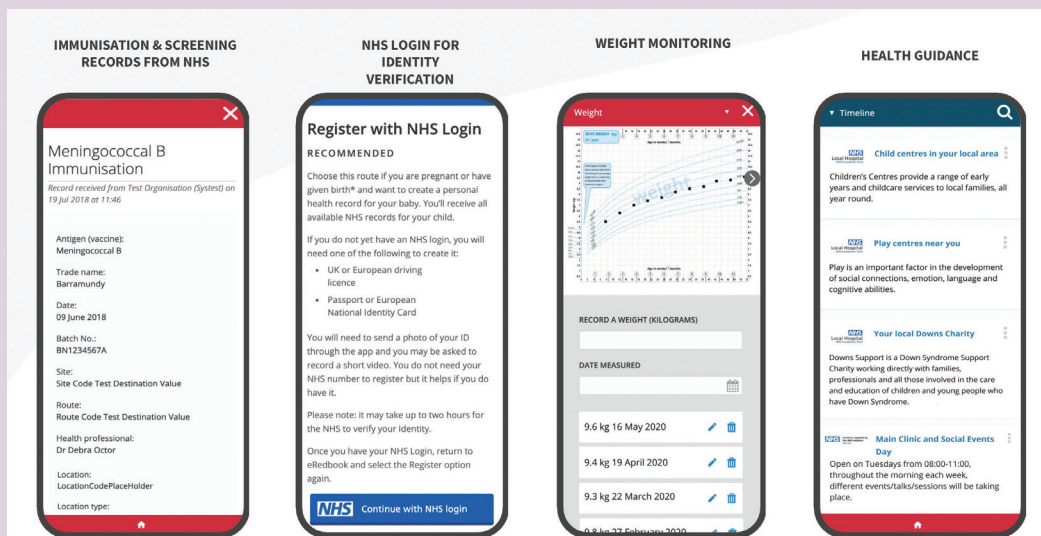
59. Green D, O'Shaughnessy J, Starks G, Sloggett R, Buchan I, Woods T. *Open Life Data to support healthy longevity for all*. *Lancet*. 2021: [http://doi.org/10.1016/S2666-7568\(21\)000081-7](http://doi.org/10.1016/S2666-7568(21)000081-7)

60. *Levelling Up Health*, <https://appg-longevity.org/levelling-up-health>

## Use Case 1

**eRedbook for Life (example).** The “eRedbook” has been successfully used to put children’s data into their parents’ hands to enable parents to digitally monitor the wellbeing of children aged 0-5 including their weight. This use case proposes to extend eRedbook benefits to the entire life of those individuals developing new tools based around life data to support weight management in adults. Sitekit would combine prototype work done to provide clinician-led weight management tools that enable the digital interaction of the clinician with the patient with new work to integrate life data obtained from 3rd party sources outside healthcare.

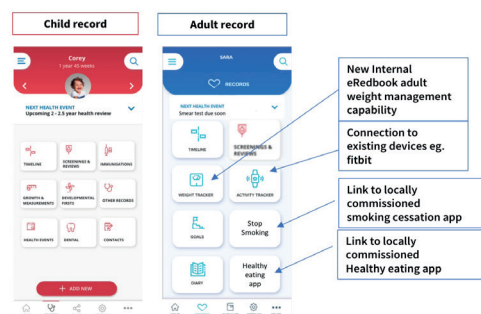
Send information and health data direct to parents without manual intervention



## eRedbook for Life – Concept example promoting well ageing

### Weight Management – Patient

- Combined offering of inbuilt capability alongside commissioned apps
  - Holistic approach for patient
  - Single point of access made simple
  - Focussed around all aspects of illness preventative action
  - Linked into fitness trackers and digital scales

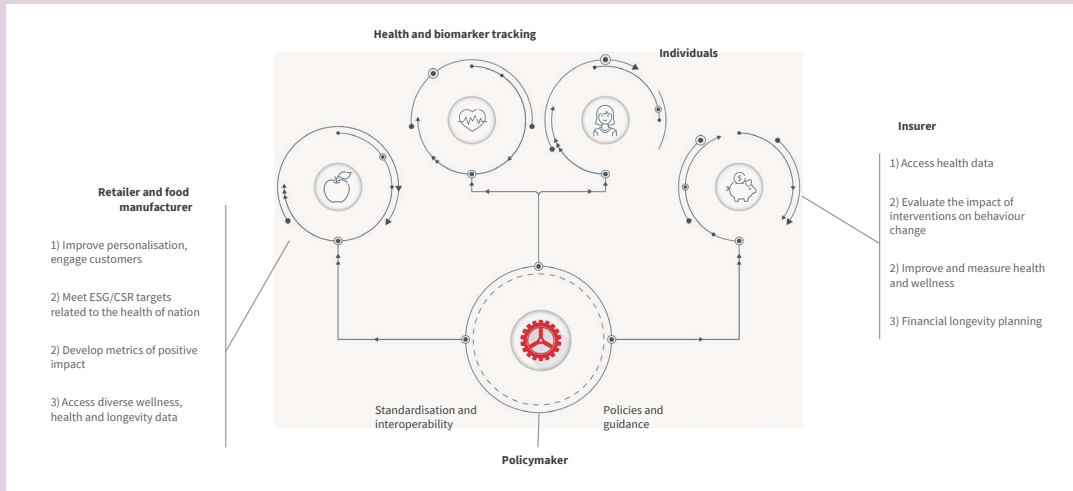


eRedbook

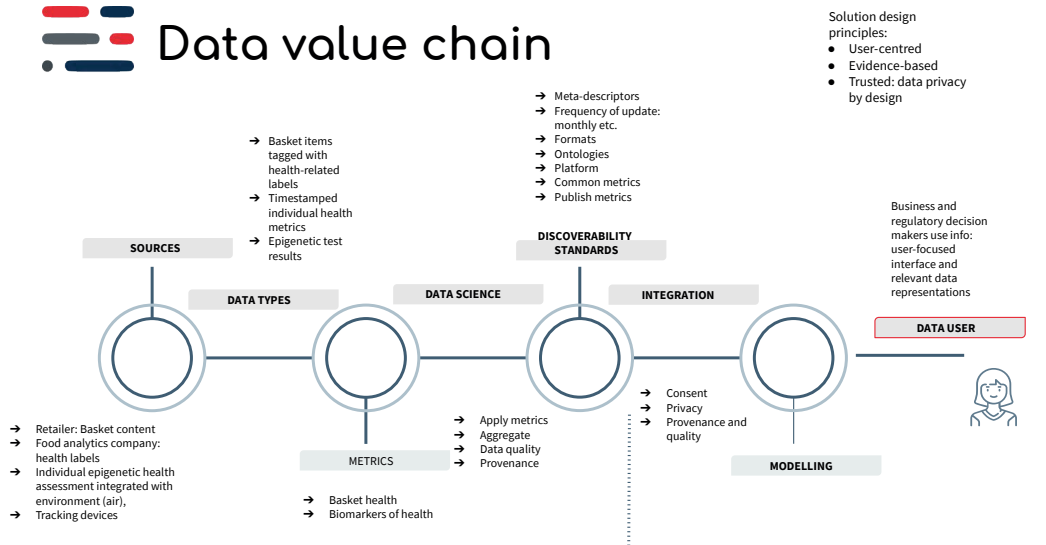
Eric Kihlstrom and Michael Catania from Sitekit are project leads. Further information is available by contacting [eric.kihlstrom@gmail.com](mailto:eric.kihlstrom@gmail.com) or [michael.catania@sitekit.co.uk](mailto:michael.catania@sitekit.co.uk).

## Use Case 2

**Ajax: The Healthy Ageing Data Exchange:** The initial use case of this project combines epigenetic and shopping data to identify and incentivise the behaviours that advance healthy longevity. It uses a federated learning architecture to securely integrate and assess datasets shared by food companies, insurers and health analytics providers to demonstrate how valuing Life Data can promote better engagement and preventive health behaviours.



## Data value chain



Svitlana Surodina and Dr Richard Siow from Ageing Research at King's College London with partners Legal & General, Muhdo, Skein Group, and Spoon Guru are on the project team. Further information is available by contacting Svitlana Surodina on [svitlana.surodina@kcl.ac.uk](mailto:svitlana.surodina@kcl.ac.uk).

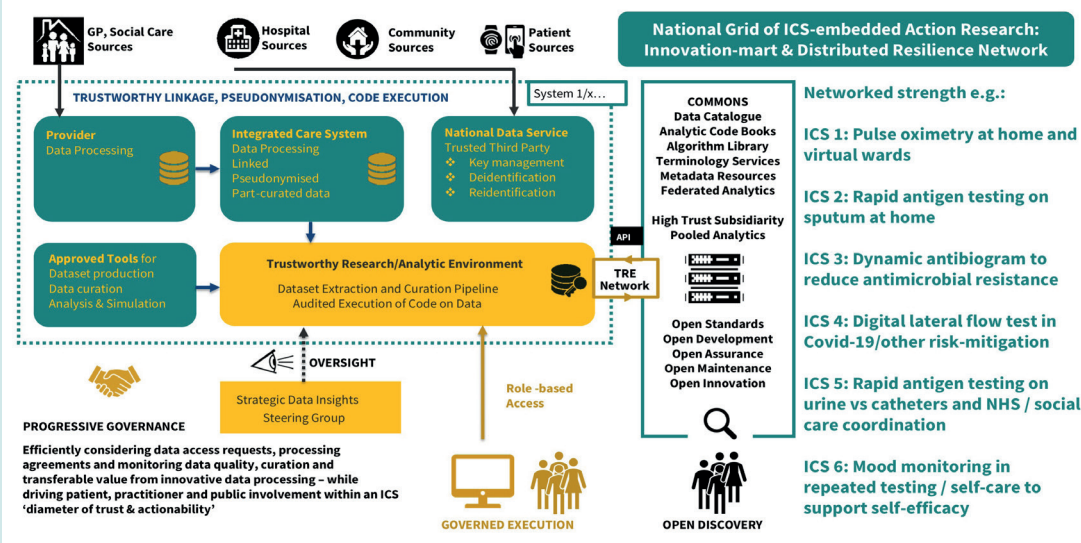


## Use Case 3

**Civic Data Cooperatives (CDC): Combined Intelligence for Population Health Action (CIPHA, [www.cipha.nhs.uk](http://www.cipha.nhs.uk))**- The CDC envisages a national grid of health action-research in civic critical masses covering NHS and Local Authority Integrated Care System (ICS) or similar footprints covering around 3 million people per system. Each regional/ICS node of the action-research grid has sufficient trustworthy data governance and processing capability to generate intelligence that improves local services and to generate insights or innovations of likely value to the wider grid. Any ICS can borrow strength from other ICS developing/validating/updating algorithms to support the decisions of patients, practitioners and population health managers – interoperability is ensured by each civic data cooperative sharing a common open life data framework. The role of the CDC is to create a 3rd space where civic and commercially generated data can be brought together, with the support of residents and patients, to generate new insights and support the development of new products. The CDC is working to create a network of CDC and TRE services that provide access to data and support for wider design challenges facing commercial companies working to build novel health solutions.

### What is CIPHA? Combined Intelligence for Population Health Action

<p><b>A cloud-based shared care record and population health service</b> built on mature infrastructure. 1m-5m citizens per instance</p> 	<p><b>Operated by 12 large care systems</b> covering 16.7m patients</p> 	<p>Combines extensive data from <b>all health and social care settings</b></p> 	<p>Currently storing <b>35bn records and over 500bn data items</b>. Expected to double this year</p> 
<p><b>Collaborative work on use cases</b> with all development available FoC to all members</p> 	<p><b>Local decision making</b> and transformation teams means local ownership of data</p> 	<p><b>Near real-time data used for direct care and population health.</b> Includes tools to action insights</p> 	<p>Same near-real-time data <b>used de-identified</b> to drive planning, quality management and research</p> 
<p>Academic research supported by <b>Trusted Research Environment (TRE) federation</b></p> 	<p><b>Managed by central NHS programme team</b>, driven by senior local leaders (mostly CEO level)</p> 	<p><b>Formal programme monitoring and evaluation</b> (incl. benefits management)</p> 	<p>Supplied by the <b>British company</b> that developed / runs National Immunisation Management service (NIMS)</p> 



Professor Iain Buchan and Dr Gary Leeming from Liverpool University are project leads. Further information is available by contacting Professor Iain Buchan on [buchan@liverpool.ac.uk](mailto:buchan@liverpool.ac.uk) or Dr Gary Leeming on [Gary.Leeming@liverpool.ac.uk](mailto:Gary.Leeming@liverpool.ac.uk).

In total ~110 people were involved in the Consultation (see Acknowledgements). Of those completing the survey in the wider consultation (N=55), 36% were health & care professionals, 34% were data professionals, 16% were policy experts; all sectors were represented with 31% in the private sector, 13% in the public sector, and 28% answering as citizens and consumers.

While most of the research was qualitative in nature, key quantitative findings are:

- The greatest challenge was cited as public trust, followed by communication
- The greatest opportunities were considered to be making health and care easier for individuals, enhancing the health and care system, managing consent to data usage and protecting individual privacy
- In terms of an overall accountability role for health-relevant data, of the options presented the greatest support was found for the NHSX and an independent non-profit as lead or co-lead, with the Department of Health, CDEI and ICO having potential role(s) in governance.

The qualitative findings have informed the Framework and the report, including recommendations.

The Confederation of British Industry (CBI) is engaging with industry to feed into the Framework's development as part of its commitment to the Business Framework for Health,<sup>61</sup> published on 18 October 2021, to measure and incentivise business contribution to health and bring 'health' into ESG investing. The CBI released the survey to its health and life science members and conducted a round table on 11 November 2021 to validate ideas and test assumptions.

### 3.5. How can we (re)build trust and public confidence?

Trust is hard to establish and easy to lose. It is imperative to continuously remember and learn from previous experiences around data sharing that have led to costly errors: such losses are monetary and reputational in the context of public trust and private sector confidence.

*Trust is an outcome of systems that are trustworthy. The framework we are proposing aims to set the standard for 'trustworthiness'.<sup>62</sup>*

#### **Trust is not a technology issue—how it is used is.**

We know from our own personal experiences of clicking 'yes' to Terms & Conditions online that we do not fully understand the implications of sharing our data. The burden of responsibility, therefore, must lie with the commercial and non-commercial actors in the system and those who regulate them, to fully address what is proportionate to both deliver health outcomes and to protect individual rights.

Making information clear to users (whether individuals or related businesses) is often underestimated as a task and, as such, our primary recommendation is that clear and usable communication be treated as a critical prerequisite and also outcome of any data sharing work. We highlight that simply having communication outputs is not the same as making them usable. The risk inherent in underestimating this task is the permanent undermining of trust in the health & care system itself, across both the commercial and public sectors.

It is also important to note that transparency and information sharing aren't the end, they are a means to an end.<sup>63</sup> Transparency must have a purpose, and it's usually not just 'reassuring people' but more to enable accountability, so that information can be shared for the purpose of scrutiny by both the public and experts.

*To have privacy, a citizen should not have to learn how the NHS works.<sup>64</sup>*

*How organisations manage personal data, both public and private, is not understood by the public, on the whole. Who 'owns' data, especially healthcare, is also not clear. Transparency regarding application and benefits of data need to be clear to allow the public to make informed choices.*

Haider Husain, COO, Healthinnova Ltd & NHS NED

62. Onora O'Neill (2018) Linking Trust to Trustworthiness, *International Journal of Philosophical Studies*, 26:2, 293-300, DOI: 10.1080/09672559.2018.1454637 63. Complete transparency, complete simplicity How can the public sector be meaningfully transparent about algorithmic decision making? *Britain Thinks*, 17 June 2021, [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/995014/Complete\\_transparency\\_\\_complete\\_simplicity\\_-\\_Accessible.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/995014/Complete_transparency__complete_simplicity_-_Accessible.pdf) 64. <https://medconfidential.org/>

### 3.6. The role of data: how can we assess its potential value?

“Value arises from data when businesses create jobs or become more productive; when governments deliver more effective public services; when our environment is clean and diverse; and when people live happier and healthier lives”<sup>65</sup>

A cohesive approach to assessing the value of data combines a number of areas:

- Strategy: what is the overall value to stakeholders and alignment with systemic digital transformation?
- Governance: who defines ownership, control, rights and responsibilities?
- Quality: is the data of sufficient potential quality for the application?
- Fit: is the data materially impactful to the application in question?
- Valuation: what value (to the desired outcomes, financially, socially) has the exchange of data?
- Management: how is data managed by stakeholders (internally, by the subject and third parties)?
- Monetisation: how should effort and value be compensated? Is there financial or non-financial reciprocity for participants?

Data increases in value the more it is connected. Whether generating financial, social or environmental impact, unlocking trusted access to data is fundamental to value creation<sup>65</sup>.

As such, data is not a ‘technology’ issue. It must be treated as part of an organisation’s business value and its operational processes, including being addressed as part of GRC (Governance, Risk and Control) processes. Such GRC processes should be made clear—and made usable—by data subjects and data users.

All organisations that manage such data must have a clear data strategy that addresses transparency, data rights, consent management and helps users across the value chain clearly understand their decisions and their implications. Investors, funders and those in procurement must mandate that such processes exist as part of their standard processes, and acknowledge that there are substantial potential liabilities if they do not.

Addressing a full risk-benefit analysis and formal evaluation<sup>66</sup> should, however, not act to inhibit innovation. There are material benefits that can be realised as a result of innovations that can better access data.

We recommend continuous collective action and collaboration—including broad stakeholder engagement—be managed through an independent body that can help to shape, define and implement rules. Such rules will span both voluntary Codes of Practice as well as formal regulations and laws.

*Formalising processes for accessing and sharing ‘health related’ data from outside of healthcare contexts is a natural but crucial next step in health data innovation. People need to see a clear purpose for the use of their data and trust the system. There is a paradox as often the most valuable data do not have a defined purpose, especially when it comes to using artificial intelligence: the more data you link the more you will get novel insights; this is difficult to translate into a very clear value proposition at the outset.*

Aidan Peppin, Senior Researcher, Ada Lovelace Institute

65. Value of Data Summary Report, 26 February 2020, [https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value\\_of\\_data\\_summary\\_report\\_26\\_Feb.pdf](https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value_of_data_summary_report_26_Feb.pdf)

66. Applying new models of data stewardship to health and care data, 16 July 2020, <https://theodi.org/article/applying-new-models-of-data-stewardship-to-health-and-care-data-report/>



*This is about purpose-led innovation, and there is an army of purpose-driven innovators in one of the world's largest workforces, the NHS, working in partnership with their local authority social care and public health colleagues in towns, cities and regions. These civic partnerships have made some profoundly progressive responses to the Covid-19 pandemic, particularly in data uses to integrate care and public health measures. We want to persist gains that we've had from working with greater agility in pandemic responses. The biggest opportunity is integration. The big question is how do you deliver the integration of care with critical mass across local government, the NHS, academia and business. How do you sustain it? And then, how do you scale?*

Professor Iain Buchan, Chair of Public Health and Clinical Informatics and Executive Dean of Institute of Population Health, University of Liverpool

*It's great to have a vision of having data combined across sectors but in reality, all that data will be held in different systems, different formats, with different levels of processing and access. To operationalise this kind of approach you need to think about common standards and interoperability to allow this data to be combined and analysed. We are involved in the Alzheimer's Disease Data Initiative (ADDI) which aims to increase interoperability of global data platforms to accelerate dementia research. How you do this in one single approach for all sectors is less clear, but clearly, the development of Trusted Research Environments and federation will be ever-more important'*

Susan Mitchell, Head of Policy, Alzheimer's Research UK

### 3.7. How can we address commercial realities and evolving models?

With the right processes and mechanisms in place, an Open Life Data Framework has the potential to transform health, moving away from the illness model to a more preventative health paradigm.

Businesses have a profound impact on social determinants of health as employers and through their social and economic impact on local areas and wider society. Emerging integrated care systems (ICSs) should regard businesses as anchor institutions alongside local authorities, local NHS and third sector groups, and will have a large role to play in improving the social determinants of health.

The Framework could:

- Stimulate open innovation to make it easier for innovators to access and share health-relevant data and harness data-intensive technologies such as AI
- Foster and greater collaboration between public, private and third sector at a time when place-based approaches and integrated care systems (ICS) are central to levelling up health
- Increase access to a wider range of more individually tailored products.
- Empower people and businesses to make more informed health and care decisions.
- Make it easier for people and businesses to access and compare health and care services.
- Support greater personal health and care management.
- Help widen access to advice and support in health and care decision-making.
- Improve competition among health, wellbeing and care service providers, spurring innovation, and development of new services and increasing demand.
- Harness Britain's strengths in genomics, life sciences, artificial intelligence and tech sector (across all verticals such as cleantech, biotech, healthtech, agotech, fintech, and insurtech) to create a significant commercial market where Britain can lead the world. contributing to the health of the UK economy.

New commercial models informed by the NHSX Centre for Improving Data Collaboration's Value Sharing Framework<sup>67</sup> are needed to refine and adapt them to suit the needs of all stakeholders in a publicly acceptable and trustworthy way. These must be tested in the real world to help understand what the potential benefits and harms, blockers and mitigations might be.

As part of this greater public awareness of the risks and benefits of sharing data with the private sector is vital. Real examples will be needed to reassure the public and the media, and to shift public opinion, especially regarding privacy. It is worth noting that Health Education England's compact with CILIP (Chartered Institute of Library and Information professionals), Libraries Connected (public libraries) and Arts Council England should help: this is a 5-year initiative for citizens to have the health literacy skills, underpinned with digital skills and resources to make shared decisions and manage their health and wellbeing.

*We've spoken to so many groups of public about this, they're very unwilling to accept commercial gain, it is all about demonstrating public benefit and public good ....they don't understand that the reason organisations are in is to make money and to commercially benefit. People just don't like that and then their minds go to 'they're going to sell it', 'we'll be targeted by marketing and insurance companies', they're going to abuse this information.*

Michelle Mackie, Head of the Qualitative Research and Engagement Centre, Ipsos Mori

67. NHSX Value Sharing Framework due for publication in 2021

*‘When the media or public voice concerns over companies profiteering from their NHS medical record data, for example, it’s not that they don’t want any profits to be made of data, it’s that they don’t want to feel exploited for this purpose and/or they want any profit to be proportionate to the value back to the NHS, patients and the public’*

Aidan Peppin, Senior Researcher, Ada Lovelace Institute

*‘The vision to combine non-health data to inform health data decision-making is one of the biggest opportunities set out in the framework and central to underpinning lots of potential across private and public sector to improve public health outcomes. And there is much potential for combined data to unlock innovation in the commercial sector however, this will have to be accompanied by the need to address the biggest challenges around public perception, trust and consent to ensure health data is used responsibly and transparently.’*

Susan Mitchell, Head of Policy, Alzheimer’s Research UK

### 3.8. Are there potential regulatory levers?

Transparency and access are the basis of a free-market economy. Organisations will, naturally, gravitate towards ‘gatekeeping’ to maximise the perceived value of their assets. However, in a digital-first economy, data is unlike a ‘traditional’ asset: it increases in value the more it is connected. This pushes the value chain to a service-focused model and, rather than a ‘race to the bottom’ can drive towards quality of service as a differentiator.

A greater focus on preventative health interventions and the development of a healthy longevity marketplace in a more agile open innovation ecosystem to harness data offer the UK a major opportunity to become a world leader in improving health, reducing health inequalities and especially in tackling the growing costs of an ageing population. The UK has globally excellent academic, biotech and pharma R&D expertise, pioneering SME sector, and unrivalled health data and clinical trials networks through the NHS.

Making healthy longevity a whole-of-government priority will fulfil the manifesto pledge of 5 extra years of healthy life expectancy while minimising health inequalities, and provide major health, social and economic rewards. Possible regulatory drivers to incentivise the development and commercialisation of a ‘healthy longevity marketplace’ include:

- Support ‘catalytic public investment’ to draw private money and encourage pension funds (including DC pensions to invest in VC or scale-up companies) and insurers to back healthy longevity projects, as announced in the 2021 Budget
- Provide simplified investment vehicles like Enterprise Investment Scheme and widen the scope for eligibility of R&D tax credit reliefs for emerging technologies in geroscience, digital technologies and ageing biomarkers to increase investment and support sector growth, to create a significant commercial Healthy Longevity industry where Britain can lead the world.
- Promote frictionless data sharing and trade in line with the thinking of the Taskforce on Innovation, Growth and Regulatory Reform<sup>68 69</sup> and UK Innovation Strategy.<sup>70</sup>

Other regulatory levers include developments in AI that can bring a wide array of economic and societal benefits across the entire spectrum of industries and social activities. For example, the UK requires an equivalent of proposed EU regulation on artificial intelligence.<sup>71</sup> Meanwhile, the Care Quality Commission (CQC) is starting to look into how AI can be used within social care.<sup>72</sup>

68. UK unveils post-Brexit global data plans to boost growth, increase trade and improve healthcare, 26 August 2021

<https://www.gov.uk/government/news/uk-unveils-post-brexit-global-data-plans-to-boost-growth-increase-trade-and-improve-healthcare>

69. Taskforce on Innovation, Growth and Regulatory Reform independent report, 16 June 2021

<https://www.gov.uk/government/publications/taskforce-on-innovation-growth-and-regulatory-reform-independent-report>

70. UK Innovation Strategy; Leading the future by creating it, July 2021

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1009577/uk-innovation-strategy.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1009577/uk-innovation-strategy.pdf)

71. Laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, Regulation of the European Parliament and of the Council, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>

72. The Health Foundation Data Project. Strengthening Social Care Analytics.

## 4. UNLOCKING VALUE

### 4.1. What are the priority considerations for value creation?

#### 4.1.1. For people (citizens, patients)

Research conducted by the National Data Guardian (NDG), Understanding Patient Data,<sup>73</sup> NIHR and other groups show that the public supports the idea of their data being shared, to benefit not just themselves but also for the wider benefit of the public. The recent report by the NDG, Putting good into practice,<sup>74</sup> highlighted that the public did not feel that data had to be used only for its original purpose if it was bringing wider public benefit. Their main priority was that the data is used in a relevant way, with the legal and ethical safeguards in place.

The Framework will enable further innovation to engage people in their health, to ‘get citizens closer to their data’ as expressed in the draft Health and Social Care Data Strategy. The pandemic has greatly accelerated digital innovation, and almost 2 million people are now using the NHSApp to access NHS services. Having health and wellness increasingly accessible via people’s smartphones will further expedite the behavioural and cultural changes underway, but this should not be at the expense of those who are not as digitally literate or enabled.

We hope this work can aid understanding of how data can be harnessed to improve health outcomes and reduce health inequalities.

*‘It is a concerning situation that members of the public are more willing to hand over data to social media companies than the NHS. The public requires a greater degree of assurance and transparency about the way the NHS handles their data. This issue can be addressed via effective educational means.’*

Louis Holmes, Senior Policy Officer, Care England

*‘I see public understanding and trust as the major barrier. I don’t think regulation is as important as communication and trust (which depends on great security!)’*

Eric Kihlstrom, Chairman, Centaur Robotics

#### 4.1.2. For the public sector

We have a chronic disease epidemic, much of which is driven by social, economic and commercial determinants of health. We need to reduce demand on our NHS and care system and focus more on addressing the root causes of ill health using a system change approach. The Framework will drive awareness of the need to look at health-relevant data beyond that collected by the NHS and care system. It will support the development of Integrated Care Systems (ICSs), help the new Office for Health Improvement and Disparities with preventative health strategies and complement the ONS Health Index plans to develop health as an asset across three key areas: healthy people, healthy lives and healthy places.

73. Quick guide to explaining how patient data is used, Understanding Patient Data, October 2021, <https://understandingpatientdata.org.uk/sites/default/files/2021-10/Resource%20quick%20guide%20v%202.0.pdf>

74. Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data, National Data Guardian, 14 April 2021, <https://www.gov.uk/government/publications/putting-good-into-practice-a-public-dialogue-on-making-public-benefit-assessments-when-using-health-and-care-data>



### 4.1.3. For the private sector

The Open Life Data Framework will make it easier to maintain trust with users by firmly addressing their data rights, while also unlocking the ability to connect datasets for innovation. The aim is to help develop foundations for businesses of all sizes, including non-profits and social enterprises (for example, community interest companies limited by guarantee) and the start-up sector (which finds it very challenging to develop and scale in the NHS).

With robust processes and systems around data sharing, consent management and governance, it aims to catalyse collaboration between the public and private sectors, in the interest of maximising value for public benefit, addressing the needs and interests of taxpayers, government and businesses.

The Framework will reinforce the role that business can play in population health, and help to identify and share private sector data that matters most to keep us healthy. With equal measure, it will reinforce the rights of the individual, reduce potential harms and ensure privacy and wellbeing is not undermined.

The Framework will feed into the work of Business for Health (another recommendation from the APPG for Longevity's Health of the Nation report) that published *Business Framework for Health: Supporting businesses and employers in their role to enhance and level up the health of the nation*<sup>75</sup> and accompanying Index for Health in three key areas:

1. Direct impact: Business influence on employee health
2. Secondary impact: Business influence on health via products and services sold
3. The external influence of business on the communities in which they operate and the wider environment/society

The Business Index is a prelude to a wider piece of work to get 'Health' into Environmental, Social and Governance (ESG) frameworks - that is, ESG. Health is where the climate change agenda was 10 years ago, and now is the time to guide more investment and innovation into health guided by ESG mandates as we do for climate change, applying them to healthy life expectancy and societal health.

### 4.1.4. For the third sector

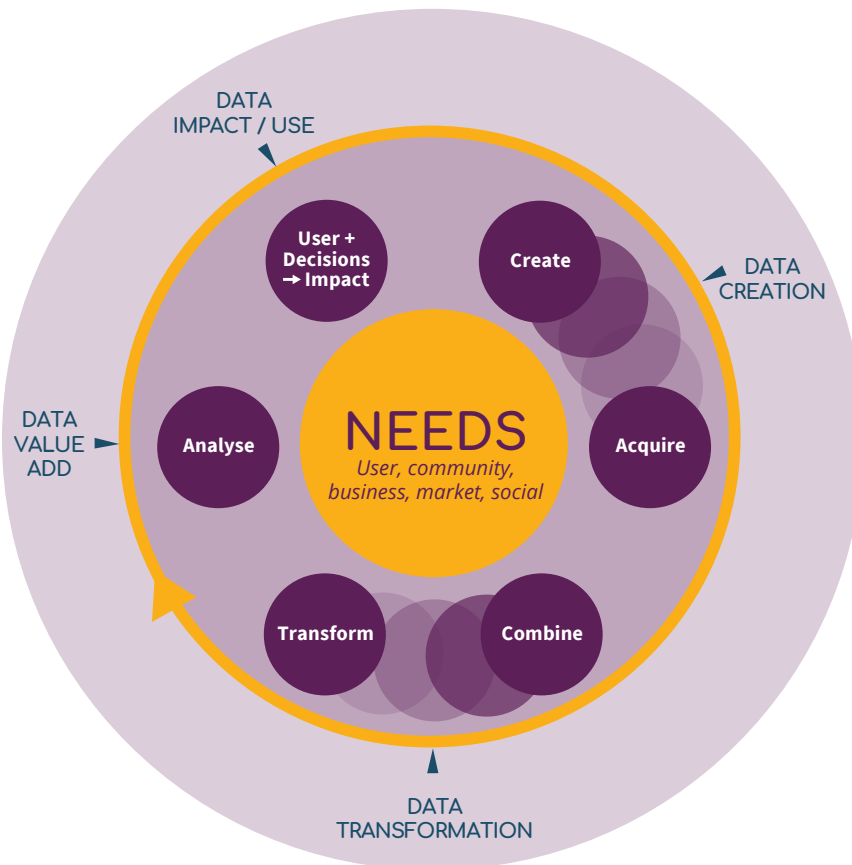
The framework will enable the opening up of access to data for third sector organisations including anchor institutions, and scale-up place-based innovation and local public health intelligence for the benefit of communities and wider society. This is of particular relevance in the context of Integrated Care Systems.

Again, managing the rights of the individual is at the heart of the framework, building a trust framework for the health ecosystem as a whole that can make it easier for participants to engage with clearly defined principles and practices.

75. *Business Framework for Health: Supporting businesses and employers in their role to enhance and level up health of the nation*, <https://www.businessforhealth.org/new-events>

## 4.2. How can we understand the data value chain and consent?

The value exchange between the consumer and the service provider or supplier must be clearly communicated and understood.



28

To understand what is possible, what opportunities and risks exist, and what design parameters should be considered, it is important to map out the **data value chain**. Data sharing consent spans three domains: regulation, consumer and market.

**Regulation:** Rules that enable an open market and address the needs of the market.

**Consumer:** Rules that enable consumers to do things, and that protect them.

**Market:** Rules that enable companies to do things and that protect them.

Regulatory approaches can mandate consistency for consent, address the challenges of onward data sharing and strengthen the principles of GDPR, its application and enforcement.

The responsibility to put in place a trust framework for data sharing lies with the market and regulatory environment. The consumer (individual or group) cannot and should not be expected to understand the rules and regulations in order to use the system. However, there should be sufficient transparency, rights, protections, processes and controls in the system to make it clear that people have protections, have rights and have clear modes of redress. Together these can form a basis of trust.

Equally, to make clear and visible why, what and how data is shared it must be tested with real-world examples that can be understood by a wide audience.

To create a trustworthy and trusted data sharing culture, we need to understand many areas:

- which actors exist in the chain (regulation, consumer and market)
- what data exists
- where it is coming from
- where it is going to
- how it will be shared
- how it will be used
- who is defining the rules and overseeing equitable governance
- what modes of consent and ongoing consent management exist within the system
- where liability sits and how it is transferred
- how security is implemented and enforced
- what modes of redress exist to facilitate remedies
- how easy is it for users (actors) to have an understanding of the system (e.g. is consumer consent ever truly 'informed', do companies understand and know how to apply the rules?)
- how usable is the system—can users act with confidence (e.g. modes of redress)

The data value chain can be viewed as a continuous cycle whereby information and insight is created and/or acquired then combined or transformed into a structure that can be analysed. Analysis adds value in the form of insights that can be used to create impact. The outcomes from the usage of data can then be fed back into the process to manifest iterative learning and development. This is a continuous process that can operate at different scales.

The cycle may take place through the work of an individual, a team or a project within an institution; equally, it could describe the activities of a global collaboration of multidisciplinary teams operating a virtual relay to add value to an initial dataset. As such, there may be different users for each cycle (e.g. the users of the initial dataset adding value to it and becoming the providers of a new dataset).

There is no clear separation of 'data' and 'service' in this case: whether the output is 'data', 'information' or 'knowledge', it is an identifiable output that can be represented digitally in an interoperable format with clear licensing terms. The reliable provision of these digital outputs (in terms of quality, timeliness and the rules of the system that govern usage) to the next actor in the supply chain constitutes a service. 'Data' should therefore not be viewed as an 'asset' (akin to a 'gold') that sits within a neat boundary but as a continuous stream of information that can inform action.

As sensors, applications and the complexity of computing increase, this process will be increasingly automated and hard to understand. Organisational processes must address the interaction between humans and machines to both understand transparency, to maximise impact and to minimise harm.

All of the parameters highlighted above are critical to the creation of rules. We illustrate this with the examples below, through the lens of the individual:

1. an individual wishes to give permission to their insurer to access their specific personal, formal medical records for a time-limited period.
2. an individual wishes their formal medical records to be used solely for the purposes of their direct care and treatment
3. an individual wishes to share a set of formal and informal (e.g. a health app) data with a commercial service to help them manage their wellbeing.

## 4.2.1. What is informed consent and consent management?

It is imperative that transparency be mandated throughout any such framework. Trustworthy consent and consent management is dependent on both the transparency as well as the usability of any system: to be able to act upon how data may be used. To enable consent requires clear definitions of rights: who has permission to control consent, when, how, and for what purposes.

There will exist different rules to describe consenting to the **transfer of data** and the lawful basis for **processing that data** under GDPR. For example, a firm may still derive some form of value from data a consumer has previously shared using a consent period even if that consumer has let such consent lapse (retroactive deletion may require a further request, or the implementation of market-wide rules to enforce data minimisation through the lapsing of retention rights).



As such, consumer primacy must lead our thinking. Can consumers have clear control around data sharing as a process (as distinct from data processing)? Can the value exchange between the consumer and the company be made clear in a manner that is acceptable to both parties? Can redress be affected in a reasonable manner?

### GDPR

- GDPR applies to any use of personal data, or any aggregated data through which it is possible to identify a living person (e.g. through dataset combination).
- Data protected by GDPR generally requires opt-in consent to access, unless suitably aggregated or otherwise anonymised.
- Aside from obtaining individual consent, the GDPR contains certain provisions for accessing protected data if ‘substantial public interest’ can be demonstrated. We explore some of these issues in this report. ‘Substantial public interest’ is not in itself defined in either the GDPR or UK Data Protection Act. A summary of conditions [4] is provided by the Information Commissioner’s Office (ICO).

## 4.2.2. What is data governance?

Data governance is currently fragmented and this is unlikely to change. To address the broad, cross-sector, decentralised and distributed nature of the data value chain will require interoperability and cohesion between sectors.

There are many existing legislative and regulatory mechanisms: a use-case-led approach to review and test a cross-sector approach could help yield, at a minimum, a consistent approach to terminology and definitions and, ideally, to a common approach for consent management, consumer experience and interoperability.



BEIS Smart Data Review, 2019 define Smart Data as follows:

“Smart Data enables consumers, if they wish, to simply and securely share their data with third parties, to enable them to provide innovative services. The UK’s data protection laws already give consumers the right to request that businesses provide their data to Third Party Providers (TPPs) in a commonly used format - this is known as the right to data portability. ‘Smart Data’ represents an extension of this right and provides an enhanced framework for sharing consumer data that allows for further innovation.”

A Smart Data strategy may provide a route to the development of common standards for cohesion and interoperability that requires organisations to:

1. put the interests of consumers first;
2. to require the provision of consent management tools that enable the implementation GDPR in a consumer-accessible manner;
3. require clear management of onward data sharing;
4. provide consumer-accessible modes of redress.

Data sharing already happens across sectors. The risk of harm occurs across the spectrum of activities from organisations’ own stewardship of data, to its use by third parties and other parties in the data value chain, including the government.

Opening up access to shared data is not an end in itself. The aim is to enable consumers to get value from their data by sharing it with organisations. New risks occur when organisations access data, store it and use it in the provision of services. Further, existing harms are simply likely to be exacerbated by the increased intelligence afforded by the data. However, old and new risks are not discrete but often interlinked, exacerbating risk and harm further. Illustrative risks include:<sup>76</sup>

### Examples of key risks not fully addressed by regulation

Access to shared data creates new operational & execution risks for firms	Use of data creates market and systemic risks	Wide sharing a consumer harm data exacerbates and risks to society
<p>Access to data: is data easily accessible, accurate and up to date?</p> <p>IT systems: Do firms have the right systems in place to facilitate new tech? Are they resilient? Is investment proportionate?</p> <p>Performance of technology: is the tech working, reliable and consistently available?</p> <p>Capabilities afforded: is data minimisation possible?</p> <p>Secure transfer of data: is it consistently secure in transit? Could data be intercepted in transit?</p> <p>Is data encrypted? Is data secure at all ends of the data chain? Testing facilities: do firms have access to facilities for testing?</p>	<p>Nefarious actors: what are the incentives and routes in for illegitimate actors? How is fraud monitored and guarded against?</p> <p>What is the likelihood of cyber attack?</p> <p>Product design features: are there tools to control data effectively? Can changes be enacted through the chain? Is there too little or too much friction?</p> <p>Data driven automation: is it working as intended? Is it tipping the market?</p> <p>Misconduct and incompetence: do firms follow the rules and standards? Do they use the data for purposes other than the consumer agreed to?</p> <p>Market discipline: what reporting and accountability requirements are there? Is there sufficient oversight, supervision and enforcement? Are there effective codes of conduct?</p>	<p>Value exchange: how should data affect the pricing of the product? Are there adequate price signals? Is the data value exchange clear?</p> <p>Social scoring: do firms or government create scores for people or businesses which reduce access, choice, social mobility, aspiration or democracy?</p> <p>Discrimination: is there negative discrimination which leads to bad outcomes for some groups of people?</p> <p>Price optimisation: does personalisation lead to an increased loyalty penalty? Are consumer behaviours exploited?</p> <p>Communications: are people and businesses aware of their rights and responsibilities?</p> <p>Incentivisation, distribution and point of sale: are consumers forced to share data? Are staff incentivised to make people share? Unfair teasers, commission bias, hollowing out?</p>

76. BEIS: Smart Data Research, Report: Liability <https://gov.uk/government/publications/smart-data-research-on-consent-liability-and-authentication>

To fully develop cross-sector data sharing, sector regulators will require concurrent and parallel powers for data protection. A National Data Coordinator or Ombudsman for complex, cross-sector data sharing cases will likely be required to address the complexity of the emerging data ecosystems.

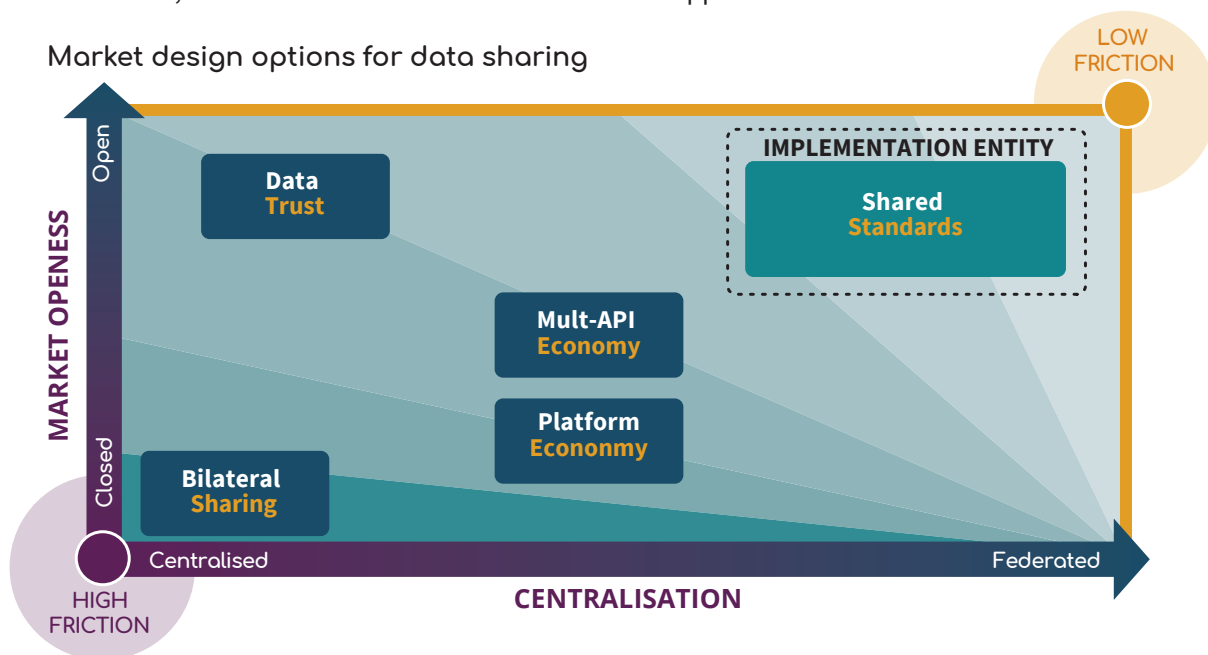
### 4.2.3. Can we help create a safe, secure open market?

In developing a functional, scalable and sustainable marketplace for data sharing, we have many models to learn from. We propose the implementation of shared standards as one potential market-enabler that could maximise openness, enable multiple actors, encourage competition and unlock value across the ecosystem. Specifically:

- Standards reduce complexity, facilitate competition and improve the consumer experience.
- Security, quality of implementation and conformance are key success factors.

By shared standards, in this instance, we mean structured machine-readable access to data (e.g. via Open APIs and operational standards) with common rules and rights (e.g. pre-emptive licensing contracts for data usage). These rules can substantially reduce the operational friction to join together existing systems and data value chains. They can reduce fragmented contractual agreements and enable relatively simple data transactions. This is of specific importance to consumers, who need to trust that there are common approaches across the market.

Market design options for data sharing



The design outcomes for Shared Data are to:

- Maximise for interoperability, cohesion and usability (government, non-government, private, academic sectors);
- Reduce friction in the system and its operation (legal, technical, cultural, regulation); and
- Open markets (diversify and extend with fit-for-purpose controls in place).

Given the current and future diversity of use cases, organisations and user needs, it is important to utilise a market design that can enable the development of multiple platforms, services, applications and consumer-facing apps based on how the user might want to use (and access) the data as opposed to how (or why) it was collected.

The starting point is the development of shared standards and the subsequent classification of data (as Open, Shared or Closed) against these standards.<sup>77</sup>

<sup>77</sup> Open Standards for Data. <https://standards.theodi.org/>

*“Can we consider enhancing the role of data governance and ethics as we move forward to more open standards? We would benefit from more citizen engagement and a stronger professionalisation of data governance and data ethics within data services teams, so that there is better dialogue and consideration of the use of data, and when we link / triangulate / persist data for products and services. At present we have more of a compliance lens to IG, and we should move towards more systematic assurance within the data service teams, so that we understand the impact of data quality / integration and subsequent analysis, which will place a stronger lens on improving transparency and data quality.”*

Ming Tang, Chief Data and Analytics Officer, Data and Analytics,  
NHS England & NHS Improvement

#### 4.2.4. Can a sandbox approach aid learning and development?

An independent regulator (or non-profit) entity can provide structure and strong governance to create a learning environment where ‘inaction’ is framed as failure and success is ‘experimentation within a transparent, risk-aware control framework’. Such an innovation framework (where teams can experiment and learn in a non-operational environment before adopting a technology, approach or system) can pave the way for innovation with appropriate controls in place to minimise potential harms.

To achieve this, incentives to encourage and mandate change are required at multiple levels:

- Personal (e.g. individuals within an organisation who lack a safe framework);
- Organisations (e.g. companies who need help prioritising data issues, care providers should get feedback if they supply data for capacity trackers, for example); and
- Markets (e.g. regulatory and legal levers of change)

Learning from the finance sector, the governance components can include:

- A modular technical architecture (e.g. add open API-based modules to existing systems rather than re-engineering underlying systems);
- Directory - a central control structure through which conditional access control is authorised to the data marketplace (nb: this does not host any of the underlying data, nor does it act as a broker. Rather it registers and tests which organisations have permission to access the market, for what purposes and under which rules);
- Legal and operational processes that facilitate data-sharing against use-cases;
- Auditable compliance with data-sharing standards – both generic and sector-specific;
- Requirements for interoperability and cohesion (including technical and user experience);
- Timeline for evolution into a data-sharing solution that meets user requirements at-scale; and
- Processes that make clear the priorities of users, the market and government.

It is worth noting that regulatory sandboxing work in the health and care sector is not new; for example, there has been interesting digital triage innovation work conducted by the CQC.<sup>78</sup>

78. CQC Regulatory Sandbox, 23 January 2020, <https://carequalitycomm.medium.com/our-regulatory-sandbox-a7b2c57bbef1>

**Mental Health & Wellbeing Sandbox with HSBC, TogetherAll, Facebook. With advisors from UK Health Security Agency (previously Public Health England), DCMS, ICO & CDEI**

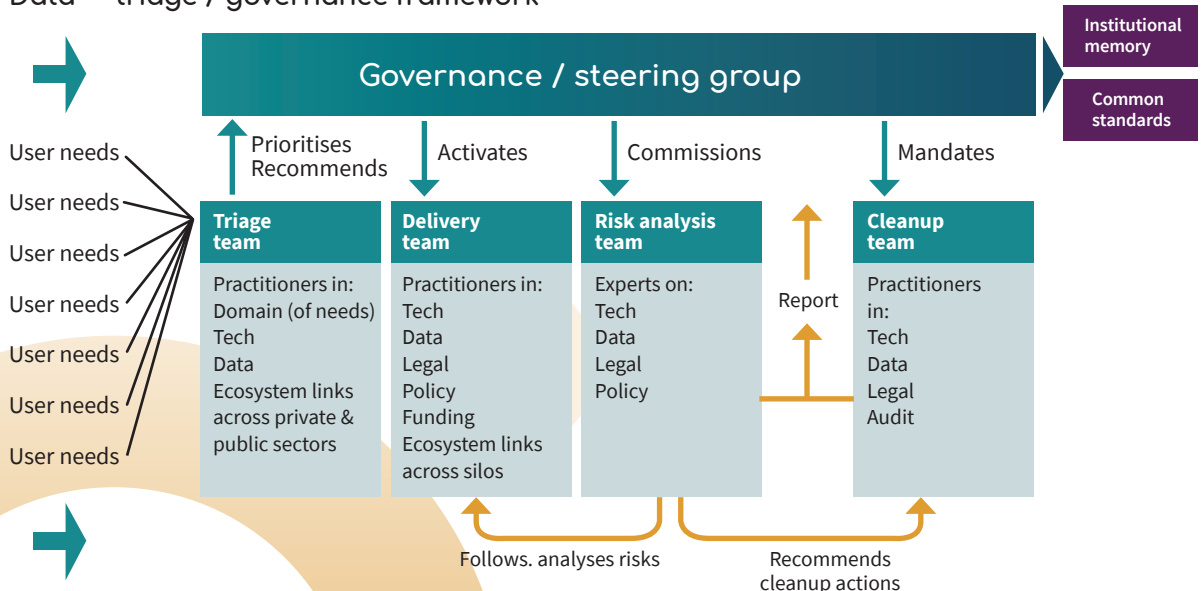
A sandbox designed and facilitated by Ctrl-Shift to develop a digital tool which enables the individual to track and proactively self-manage their mental wellbeing. It is driven by personal behaviour data, under the control of the individual, from a range of everyday digital services. The service helps the individual track their everyday activities and spot patterns to feel more in control of and proactively manage their mental wellbeing.

Mental wellbeing was chosen as the focus of this sandbox because it is rich with opportunity and potential value arising from unlocking Data Mobility, and also because this value derives from highly sensitive data that could be easily exploited to the detriment of the individual. Part of the job of good design is to simplify complex decisions and reduce the cognitive load of the user to an appropriate level while still empowering them. We have established the role of the 'Data Facilitator', a body that acts on behalf of the user with transparency and trust as a key enabler for such a service. If this is achieved, it would open wholly new areas of opportunity while preserving the freedom of the market to compete and innovate, moving in this particular case beyond the management of mental wellbeing to an opportunity for preventative healthcare at a population level.

**4.2.5. Can we triage rapid innovation while addressing risks and controls?**

Our data-enabled world will be continuously monitoring and learning. We must step up to the challenge of managing our individual agency as innovation outpaces regulation

**Data — triage / governance framework**



Triage recommends actions. Steering activates. Delivery gets on with it, quickly. Risk follows and identifies issues. Cleanup tidies up Risk and Cleanup report to Steering and generate institutional memory and can inform standards, principles and practices



In certain circumstances (e.g. rapid application development environments where fast responses are needed) a data innovation ‘triage’ process can be considered to help mitigate risk. This process aims to focus effort on ‘what really matters’ as outcomes and impacts (e.g. taking an OKR approach for outcomes and key results).

Further, recognising that there exists a very broad spectrum of data maturity across organisations we recommend that teams consider this approach in the development of data-enabled solutions. Roles and responsibilities are diverse and may need to be drawn from across the ecosystem to represent the diverse needs of different users in the value chain:

- ‘Steering’ combines expertise to ensure decisions are timely, effective and address risk (including guidance to enable regulation and innovation to keep pace with each other);
- ‘Domain’ specialists include sector or topic specialism that is based around the specific user needs;
- ‘Tech’ includes understanding everything from coding to data stores;
- ‘Data’ includes everything from data analytics and data science to algorithmic modelling;
- ‘Ecosystem links’ includes links across silos from public, private and third-party systems and from other sectors;
- ‘Legal’ includes expertise in the Domain and in data;
- ‘Policy’ includes expertise in the Domain and across relevant applications; and
- ‘Audit’ includes expertise across all areas so that the ‘clean up’ process is comprehensive.

Questions to consider include:

- What problems are we trying to solve? (based on actual user needs)
- How will data help to solve these problems?
- Who should lead? Sector / challenge / opportunity?
- When are answers needed, and how frequently?
- What minimum viable demonstrators might help?
- What institutional memory, standards, principles and practices can be created as a legacy output?

A process outline (illustrative):

1. A ‘Triage team’, comprising practitioners from the various domains, can act as the first point of call and consider the different user needs in the assessment and development of programmes.
2. A ‘Governance/ Steering Group’ is assigned to approve recommendations and prioritises responses put forward by the Triage team.
3. The Governance group can initiate rapid-response from a Delivery team who focus on solutions in the knowledge that there is a process overseeing and following them to ‘tidy up’ (rapid development is messy and leaves a data footprint).
4. The Risk analysis team ‘follow’ the Delivery team, both documenting lessons learned, making sure that the output work is compliant with standards, regulation, principles and practices and highlighting areas for review and clean up. The Delivery team relies on this independent team managing risks arising or issues that materialise, enabling them to focus on delivery.
5. The Clean up team follows the outputs of the Risk team to ensure data, code, and related digital assets are secure, deleted or otherwise made compliant.
6. The Governance team have responsibility for both oversight and capturing outputs in an incremental manner to both (a) build usable institutional memory and (b) develop and input into living standards.

#### 4.2.6. What role can Open Standards play in addressing consent-based data sharing?

Better access to commercial and open data can create the potential for target-based impact, investment, products and services but this must be balanced against the specific challenges identified. Data and technology are tools, not solutions.

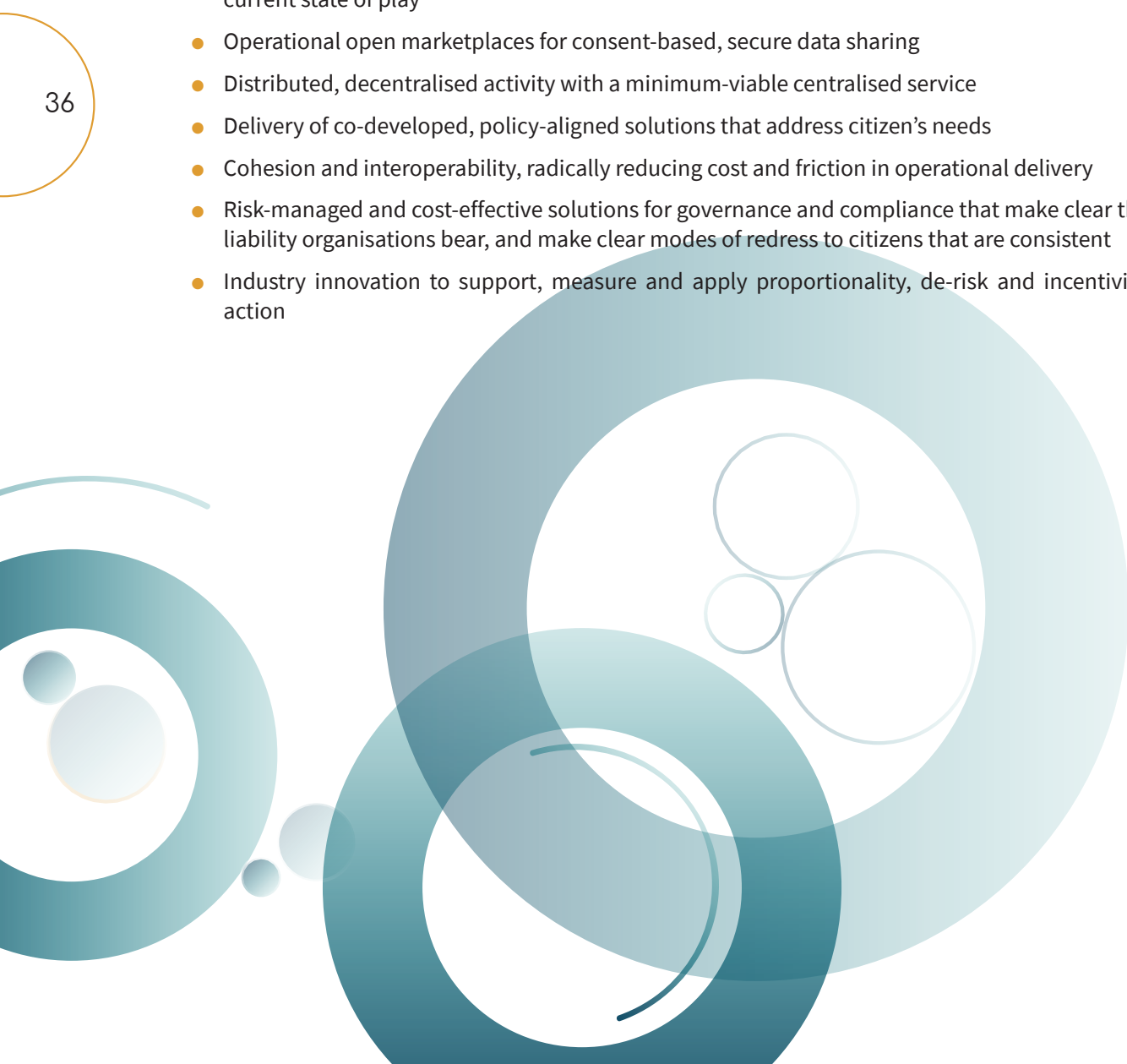
Where applicable, trusted access to data can be a critical blocker to incentives and demonstrable impact. Open standards that deliver cohesion & interoperability can unlock innovation at sector scale.

Examples of open standards that enable and operationalise data sharing include Open Banking and Open Energy. The principles and priorities for data standards, interoperability and cross-sector coordination are addressed in the National Data and FCA Open Finance Strategies.

A trust framework can inform mandates for policy and regulation, legal, operational and technical standards for interoperability. This is a systems-based approach in which disparate data can be activated and operationalised.

The approach can enable:

- Coherent and consistent consumer communication and protection that is a step up from the current state of play
- Operational open marketplaces for consent-based, secure data sharing
- Distributed, decentralised activity with a minimum-viable centralised service
- Delivery of co-developed, policy-aligned solutions that address citizen's needs
- Cohesion and interoperability, radically reducing cost and friction in operational delivery
- Risk-managed and cost-effective solutions for governance and compliance that make clear the liability organisations bear, and make clear modes of redress to citizens that are consistent
- Industry innovation to support, measure and apply proportionality, de-risk and incentivise action



### 4.3. How can we link with national priorities, policies and related programmes?

The Open Life Data Framework aims to add value to and build on the ongoing national direction and dialogue being led across NHSX. It complements the plans for health and care data while recognising the relevance and value of data outside the health and care system to enhance individual and population health and reduce health inequalities. Crucially, the Framework can answer the need for standardised integration of health-relevant data, particularly at an ICS level, to address health inequalities and aid pandemic recovery.

It supports efforts to ‘empower researchers with data to develop life-changing treatments, models of care and sights’. By increasing the scale of the datasets available to researchers and enabling more assets to be linked, while maintaining the necessary safeguards, richer data insights can be obtained.

Overall, the Framework supports the broader principles and aims of the draft Health and Social Care Data Strategy to unlock innovation at scale and harness data to move the NHS from being an illness service to one more focussed on prevention and wellbeing. This will require sharing data across government departments too- as seen with the Shielded Patients List, in which organisations including the NHS, NHS Digital, local government, adult social care, and the Ministry for Housing, Communities and Local Government (MHCLG) worked together to share data to keep people safe.

#### **Draft Health and Social Care Data Strategy**

The Health and Social Care Data Saves Strategy was published in draft in July 2021, updated in September,<sup>79</sup> and will be publishing a final version expected in 2021 (having incorporated feedback). In its draft it commits to:

- publish the first transparency statement setting out how health and care data has been used across the sector (2022)
- give citizens the ability to see what research their data has informed, and who has had access to their data, as soon as the technology allows (ongoing)
- embed an open-working approach across health and social care, consistent with protecting privacy, so the public can easily find and understand the data delivery work (ongoing)
- lead the Health and Care Information Governance Panel to develop guidance, frameworks and standards such as data sharing and transparency so the public can be confident in our handling of their data (December 2021)

79. Data Saves Lives: Reshaping Health and Social Care with Data, Department of Health, 28 July 2021 updated <https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data-draft/data-saves-lives-reshaping-health-and-social-care-with-data-draft>

### General Practice Data for Planning and Research- GDPR

NHS Digital is the national custodian for NHS health and care data in England and is responsible for standardising, collecting, analysing, publishing and sharing data and information from across the health and social care system.

The General Practice Extraction Service (GPES) is the way that NHS Digital currently extracts limited data from GP practices, to enable correct payment to GPs and to allow for some limited analysis. GPES is ten years old and is no longer fit for purpose, so a proposal has been developed to replace this service with the General Practice Data for Planning and Research (GDPR) which will reduce burden for GPs and which will be a more comprehensive dataset for research and analysis.

GP data collected by NHS Digital is coded data, for example a code which represents a diagnosis of Type 2 diabetes or a set of blood test results. It does not include free text data where your GP may have taken notes during your consultation. This would not change under GDPR.

GDPR<sup>80</sup> aims to enhance planning and commissioning of health and care services and facilitate research, including long-term impact of coronavirus on the population, healthcare inequalities, and the development of cures for serious illnesses (for example, the University of Oxford RECOVERY trial, which has found ways to improve the treatment for people with coronavirus).

NHS Digital would not collect patients' names or addresses under this proposal. Any other data that could directly identify patients (such as NHS Number, date of birth, full postcode) would be replaced with unique codes which are produced by de-identification software before the data is shared with NHS Digital. As cited in the draft Health and Social Care Data Strategy, NHS Digital will be able to use software to convert the unique codes back to data that could directly identify patients in certain circumstances, and where there is a valid legal reason. The use of this software would be strictly audited and published. This is an area that has caused concern especially amongst privacy campaigners.

The GDPR proposal was originally scheduled to start in September 2021, but due to concerns over inadequate consultation with the professional community and general public, has been delayed until further notice. NHS Digital is now working with a range of stakeholders to redesign the programme and to fully consult. No data will be collected until the National Data Guardian, British Medical Association and The Royal College of General Practitioners agree that the public has been appropriately engaged and informed and that certain technical requirements around opt-outs and the way that data can be accessed (in a Trusted Research Environment) have been met.

The direction of travel towards Trusted Research Environments (TREs) will help to alleviate concerns. The Goldacre Review is timely to provide recommendations on the best options for future activity related to TREs and data management across England, including the development of a UK-wide (and beyond) federated network of TREs.<sup>81</sup>

In addition to TREs, it is important to note the rapid development of increasingly sophisticated privacy-preserving techniques and technologies including the following referred to in the draft Health and Social Care Data Strategy:

- synthetic data: generating data that is statistically consistent with a real dataset and can replace or augment sensitive data

80. General Practice Data for Planning and Research (GDPR), <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research>

81. Response to the Goldacre Review - "Trusted Research Environments and Data Management", <https://www.hdruc.ac.uk/news/response-to-the-goldacre-review-trusted-research-environments-and-data-management/>



- federated analytics: applying data science techniques by moving code to the data, rather than the traditional approach of collecting data centrally
- homomorphic encryption: allowing computations to be performed on encrypted data
- differential privacy algorithms: enabling useful population-level insights about a dataset to be gained, while limiting what can be learned about any individual in the dataset.

The draft Health and Social Care Data Strategy commits to ‘working and coding in the open’, and states that as public money is used for the NHS, any code used to enhance health and care should be made open source to enable digital pioneers to reuse and build on this. It is important to stress that making software code ‘open’ does not mean making data ‘open’.

Importantly, the Framework will align with plans set out in the Data Saves Lives Strategy for a ‘modern architecture in which data can be accessed real-time through APIs via a national gateway’ based on the following principles:

1. All data will be validated at the point of entry to improve data quality
2. All data will be made discoverable
3. Data will not be duplicated
4. All clinical data stored will be made accessible using APIs published on the API gateway
5. People will be able to self-manage any data relating to their contact details and personal preferences
6. Organisations should be able to self-manage any data relating to them, for example, locations and types of services offered
7. Data should be digitally signed to an appropriate level

The Framework will also support and inform government plans to enable the flow of data across the system to support interoperability and develop standards that are open, maintained, usable and accessible- and which will promote frictionless data sharing and interoperability across silos (NHSX aiming to publish recommendations March 2022).

The intent of the Framework to open up innovation aligns with government plans to separate health and care data from the application- enabling innovation to take place securely and openly on top of the data (in a TRE), with applications able to access the data directly rather than rely on vendors to provide access. Overall, open standards, code, APIs and systems architecture — when applied within a clear regulatory and policy framework— can enable innovation and promote data partnerships which maximise benefits to citizens and the system.

Overall, the strategy states that the system will look increasingly to trusted research environments (TREs), currently being used by UK Biobank, ONS and OpenSAFELY,<sup>82</sup> for example. TREs allow the creation and application of models to person- and patient-identifiable data without moving the data, and provide reassurance to the public that those entrusted with their data are keeping it safe whilst still protecting the privacy of individuals.

82. OpenSAFELY, <https://www.opensafely.org/>

## Our Future Health

Our Future Health aims to be the UK's largest ever health research programme, recruiting up to five million adults that truly reflect the UK population. Public and participant trust are central to everything we do and essential if we are to succeed. So members of the public have been and will continue to be involved fully throughout the life of Our Future Health, informing our programme design and materials as part of a mature conversation about the use of health data.

Participants will be asked to provide information about their health and lifestyles and a small sample of their blood. With their informed consent, we will combine this with existing information about them, including their health records.

Combining these multiple sources of health and health-relevant information, including genetic data, will create an incredibly detailed picture of health and wellbeing that represents the whole of the UK. Researchers will use this information to make new discoveries about human health and major diseases, and drive innovation in the field of early diagnosis. Participants will also be offered personal feedback about their health if they wish to receive it. By ensuring that a diverse range of people join Our Future Health, we can make discoveries that benefit everyone and help to reduce health inequalities.

## HDRUK: Future of Federated Analytics

HDRUK aims to unite the UK's health and care data to enable discovery. Core to that objective is federated cohort discovery and federated analytics across an alliance of TREs, lifting them from purely Data Safe Havens to actionable data engines whilst providing confidence to data custodians and the public. Data access is based on an evolution of the Five Safes model, required to meet research needs for machine learning, HPC, and modern computational workflows techniques, and a Gateway currently provides a portal to datasets, tools and resources.

Although the public health data that HDRUK primarily handles is outside the scope of the Open Data Framework, the extensive knowledge within HDRUK in Federated Anonymised Analytics provides a touchstone for: data privacy preserving analysis without data movement; the necessary services (e.g. federated identity); and the complexity of data and software governance. The mantra is data visiting rather than data sharing. HDRUK's ambition is to create an ecosystem of federated analytical capabilities interoperating with the existing TRE capabilities in the UK with an emphasis on standards, machine processable metadata, and reproducible and auditable processing.

For example, the CO-CONNECT infrastructure allows researchers to safely request access to and analyse data from patient cohorts across a number of different datasets. The HDRUK Prognostic Atlas will use these federated analytical tools to perform dynamic, aggregate level queries across datasets to provide a personalised prognostic outcome, e.g. "What is a 55-year-old woman's 1-year risk of developing an inflammatory disorder if she has high C reactive protein, mild anaemia but no disease diagnosed?". Datasets from socioeconomic, administrative, environmental and commercial sources should enrich such prognostic tools.

83. *Harnessing technology for the long-term sustainability of the UK's healthcare system*, <https://www.gov.uk/government/publications/harnessing-technology-for-the-long-term-sustainability-of-the-uks-healthcare-system/harnessing-technology-for-the-long-term-sustainability-of-the-uks-healthcare-system-report#fn:5>

84. *Integration and innovation: working together to improve health and social care for all*, <https://www.gov.uk/government/publications/working-together-to-improve-health-and-social-care-for-all/integration-and-innovation-working-together-to-improve-health-and-social-care-for-all-html-version>

85. *Taskforce on Innovation, Growth and Regulatory Reform independent report*, 16 June 2021, <https://www.gov.uk/government/publications/taskforce-on-innovation-growth-and-regulatory-reform-independent-report>

## Other relevant policies

Our work builds on the report, *Harnessing technology for the long-term sustainability of the UK's healthcare system*,<sup>83</sup> published on 23 August 2021 outlining two key priorities to drive system change:



- Supporting a step-change to improve and maintain population health, centred around new ways of engaging and supporting individuals and communities, enabled by data and evidence
- Re-engineering the health system to support integrated 'pathways' for prevention, treatment and care and better outcomes for individuals.

Overall, the Framework will support the development and implementation of Integrated Care Systems, especially in relation to two forms of integration which will be underpinned by legislation in the upcoming Health and Care Bill:<sup>84</sup> integration within the NHS to remove some of the cumbersome boundaries to collaboration and to make working together an organising principle; and greater collaboration between the NHS and local government, as well as wider delivery partners, to deliver improved outcomes to health and wellbeing for local people.

The Open Life Data Framework will align with agile regulatory reform as envisioned by the Task Force on Innovation, Growth and Regulatory Reform<sup>85</sup> and support the Data Value Sharing Framework about to be published by NHSX Improving Data Collaboration Centre.

Over time, we hope that it may be developed into a Code of Practice building on the NHS's *Best Practice Guidance for Digital and Data-Driven Health Technologies*<sup>86</sup> and, potentially, elements of it regulated to better cement trust across the ecosystem of practice around this field.

## In Step with Agile Policy and Regulatory Reform

NOW 	FUTURE 
Draft Health and Care Bill	Innovation, Growth and Regulatory Reform
Draft Health and Social Care Data Strategy	Trade Bill
National Data Strategy	Trusted Research Environments (TREs)
National AI Strategy	
NHSX AI Strategy	
NHSX CIDC Value Sharing Framework	

<sup>86</sup> *Best Practice Guidance for Digital and Data-Driven Health Technologies* <http://code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology>

## 5. USE CASES - SIGNPOSTING THE FUTURE

The intention of the Open Life Data Framework is to stimulate and foster the creation and commercialisation of limitless use cases to enhance and level up health. The Framework should also make it easier for entrepreneurs and innovators developing products and services already.

One possible area of opportunity, which the Framework could support, is mental health. Here, for example, privately-held health relevant data – if the fundamental data stewardship and ethical challenges are addressed – has the potential to support a better understanding of common mental health disorders, including amongst younger people. For example, health-relevant data could help either to identify individuals/groups at risk of common mental disorders who have not been identified by health services, or to further develop our understanding of the drivers of mental health.

This, in turn, could provide opportunities to support the development, implementation and evaluation of health interventions and other innovations aimed at improving mental health outcomes. Expanded access to datasets in different contexts and populations, especially those not in contact with healthcare services, will help shed light on what drives clinical impact, or the interplay between biological, psychological, and social factors in anxiety, depression, and psychosis<sup>87</sup>.

### In conversation with an entrepreneur

*...On data, impact comes from three things: 1) longitudinal data on the same subjects; 2) putting that data in an environment where the trust level doesn't need to be high but done in a privacy preserving way, for example, in a federated manner with differential privacy; and 3) to aid machine learning and data analysis on this data set, in a way that you don't need vetting on who gets access to it...*

*...if you make the goal to allow everybody to be able to use the data, then you're going to get 1000 more great ideas than you would have if you restricted this data to a small circle managed by data custodians...*

*The move towards TREs makes sense as long as a lot of the data is accessible... it doesn't need to be in the same place but it needs to be at least accessible in a federated manner meaning where you can do a query. Gatekeeping should be minimal, but you could have a multilevel trust system; in which there are trust settings that allow you to have more access to it.*

*You could have a vetted insider layer, but you want to move towards having more of the data available for less trusted individuals... you want to head in that direction. That can be an iterative process but it needs to be the intention from the beginning, as opposed to something that gets added later on.*

*Let's not obsess over the quality of data- let's just make it available! Often data quality is used as an excuse for inaction! Our main challenge with Humanity is getting access to reference data sets and diversity of people...*

*We will achieve 100 times the impact the second you can have an open API. For example, the prevalence of Google Maps on the internet is because you can just go and grab an API and throw it into whatever your app is, there is no vetting etc*

87. Miranda Marcus, How can we balance equity and empowerment with open mental health science? 9 November 2021, <https://mirandamarcus.medium.com/?p=b5a12a2212b5>

*The NHS, if done right could be a beacon for data sharing and create billions of dollars in good business and returns back to taxpayers - making people healthy is good business and so much can come out of that data to create multiple businesses or just make the businesses that exists, better...we need open source licensing and you don't want to limit people's ability to make money.*

#### **In conversation with a life sciences pioneer**

*Janssen is already collaborating on several projects, in which the UK is very much involved, to increase the access to, and quality and usability of, real-world data for healthcare decision-making, including through the EU Innovative Medicines Initiative. Projects such as HARMONY and PIONEER define outcome measures and utilize big and deep data sources to improve patient care. Moreover, Janssen is leading the European Health Data and Evidence Network (EHDEN) project, where we are now working with ~450 million Electronic Health Records to create a large-scale federated network for real-world data research in Europe.*

Dr Nigel Hughes, Scientific Director, Observational Health Data Analytics/  
Epidemiology, Janssen Research and Development

#### **In conversation with a citizen empowerment advocate**

*We are acting as a safe data logistics hub-managing the journey through treatment, whereby all the data that is generated is copied and goes into the individual personal data store, and then they can share that data, the relevant data with relevant service providers, as and when necessary and at their choosing. We're making the individual the point of integration of their data, so that they can use their data for whatever purposes and tasks come up as they go through this journey.*

*In our Macmillan cancer work, it typically takes about 12 hours of work per individual for a link worker to manage that person through the process. We're reducing it to one hour, because we're enabling data sharing and also dealing with all the interoperability issues*

*In our housing work with Blackwood Homes, we're linking smart homes and smart sensors, and the individual's data relating to what they're doing in their home and also wearables*

*As we gather more and more data in the individual's personal data store, we're creating a data infrastructure and a new integrated data asset, which has got all health-relevant data, including from medical records but also outside the health and care system.*

Alan Mitchell, Chairman, Mydex CIC



## 6. CONCLUSIONS AND NEXT STEPS

The APPG for Longevity is grateful for the contributions from a wide range of organisations and experts listed in the Acknowledgements.

We plan to work with many organisations in the next stage of the development of this Framework, including the implementation of the recommendations in this report.

We are seeking funding and other in-kind support for the next stage, including testing assumptions of the Framework in the real-world, on use cases, sandboxes and pilot projects.

If anyone is interested in working with us, please contact the APPG Secretariat on [secretariat.manager@appg-longevity.org](mailto:secretariat.manager@appg-longevity.org)

# 7. ACKNOWLEDGEMENTS

## Authors

- Gavin Starks, Founder & CEO, Dgen
- Tina Woods, Founder & CEO, Collider Health; Director, APPG for Longevity

## Chair of Working Group

- Lord James O’Shaughnessy, Former Parliamentary Under Secretary for Health, Visiting Professor, Institute for Global Health Innovation, Imperial College; Senior Partner, Newmarket Strategy; Chair, Science, Genomics and Technology Board, APPG for Longevity

## Funding

This project was supported by the Health Foundation. The Health Foundation is an independent charity committed to bringing about better health and health care for people in the UK.

## Contributors and open feedback

The development of this report included an open consultation phase that listed out, and solicited feedback as a ‘critical friend’ on a variety of topics and frequently asked questions.

Over 100 individuals responded from across the sector. Their responses must not be considered as an endorsement of the recommendations of this report.

We wish to thank all those who fed back—their input has directly contributed to its development and recommendations.

- Nur Alkhateeb, CEO, Ipad Digital CIC
- Lisa Allen, Head of Consultancy for Data Programmes, ODI
- Liz Ashall-Payne, Founder & CEO, Orcha
- Ryan Avison, Head of Office, National Data Guardian
- Professor Richard Barker OBE, Director, New Medicine Partners, Metadvice
- Yael Benvenisti, CEO, Mediterranean Towers Ventures
- John Blakesley, Deputy Chief Executive at Milton Keynes Hospital NHS Foundation Trust
- Claire Bloomfield, Deputy Director for the Value of Data programme, NHSX
- Jennifer Boon, Deputy Director, Data Policy, NHSX, Department of Health and Social Care
- Dr Danny Bosch, Head of Commercialisation, Innovation Programmes and Intellectual Property, NIHR Central Commissioning Facility (CCF)
- Jennifer Boon, Deputy Director, Data Policy and Digital Oversight, NHSX
- Phil Booth, Coordinator, MedConfidential
- Professor Peter Bradley, Director of Health Intelligence/Chief Information Officer, Public Health England
- Liz Brandt, Co-founder & CEO and Co-Founder, Ctrl Shift
- Professor Iain E. Buchan, Chair of Public Health and Clinical Informatics and Executive Dean of Institute of Population Health, University of Liverpool
- Philip Burgess, Philip Burgess Consulting Limited
- Rt Hon Paul Burstow, Independent Chair, Hertfordshire and West Essex Integrated Care System
- Sarah Cadman, Director for Talent & Training, HDRUK
- Sam Cannicott, Deputy Director (Projects), Centre for Data Ethics & Innovation (CDEI)
- Margaret Charleroy, Head of Strategy, NHSX
- Tony Clack, CEO, The Knowledge, Skills and Experience Foundation
- Heather Cook, Head of Strategic Partnerships, Orcha
- Professor Lynne Corner, Director of Engagement, VOICE, National Innovation Centre for Ageing
- Helen Crampin, Innovation and Technology Lead, Healthy Ageing, UK Research & Innovation

- J Cromack, Co-founder, MyLife Digital and Lead Product Evangelist at DataGuard
- Dr João Pedro de Magalhães, Aging expert, Microbiologist, University of Liverpool
- Sarah Deeny, Assistant Director, Data analytics team, Health Foundation
- Dr Rupert Dunbar-Rees, Founder & CEO, Outcomes Based Healthcare
- Marc Farr, Founder, Beautiful Information
- Lord Geoffrey Filkin CBE, Chair, Strategic Advisory Board, APPG for Longevity
- John Fitzpatrick, Head of Smart Data and Economic Regulation, BEIS
- James Flint, Co-Founder & CEO, Hospify Limited
- Mr Michael Geer, Co-Founder & CSO, Humanity
- Dr Saira Ghafur, Co-Founder, Prova Health; Lead for Digital Health, Institute of Global Health Innovation, Imperial College
- Professor Carole Goble CBE, Professor of Computer Science, University of Manchester
- Nuno Godinho, Partner CPO & CTO at Empericus and Founder & CEO at Diisruptance
- Dr Richard Gomer, Lecturer in Computer Science, University of Southampton
- Dr Joanne M. Hackett, Head of genomic and precision medicine, IQVIA
- Steve Halliday, Deputy Director Digital Transformation, NHS E/I Midlands
- Fiona Hamill, Snr Government Affairs & Policy Manager, Janssen
- Jake Harrison, Policy Adviser, Health and Life Sciences, Confederation for British Industry (CBI)
- Eleonora Harwich, Director of Collaborations, NHSX
- Louis Holmes, Senior Policy Officer, Care England
- Janet Howd, The Voice Practice
- Nigel Hughes, Scientific Director, Janssen Research and Development
- Haider Husain, COO & NHS NED, Healthinnova LTD
- Joe Rafferty, Chief Executive, Mersey Care NHS Foundation Trust
- Stephen Johnston, Co-Founder, Aging 2.0
- Josh Keith, Senior Fellow, Health Foundation
- Eric Kihlstrom, Chairman, Centaur Robotics, UK Chapter Lead, Aging 2.0
- James Kingston, VP Research Partnerships, Dataswift
- Sue Lacey Bryant, National Lead for NHS Knowledge and Library Services, Innovation, Digital and Transformation Directorate, Health Education England
- Dr Amanda Lamb, Programme Director, Health Innovation Liverpool
- Dr Gary Leeming, Director, Civic Data Cooperative, University of Liverpool
- Mavis Machirori, Senior Researcher, Health Data and Covid19 Technologies, Ada Lovelace Institute
- Michelle Mackie, Head of the Qualitative Research and Engagement Centre, Ipsos Mori
- Maxine Mackintosh, Programme Lead on Diverse Data, Genomics England
- Dr Mahiben Maruthappu, Founder & CEO, Cera care
- Andrew McCulloch, Multiple NED and Chair roles in health and social care
- Dr Ian McKay, Innovation Lead-Advanced Therapies, UKRI
- Grace Melvin, Digital and Data Policy Officer, AMRC
- Dr Alison Metcalfe, Head of Professional and Clinical Services, NHS Business Services Authority
- Alan Mitchell, Chairman, Mydex CIC
- Susan Mitchell, Head of Policy, Alzheimer's Research UK
- Jessica Montgomery, Executive Director, Accelerate Science, University of Cambridge
- Tim Morgan, COO, Service Robotics Ltd
- Jessica Rose Morley, Policy Lead the DataLab, University of Oxford
- Dr Lorraine Morley, Researcher, consultant and advisor on age tech and healthy ageing innovation, Age Tech Intelligence
- Lisa Murphy, Programme Manager, NHSX CIDC
- Ms Annemarie Naylor MBE, Common Futures
- Rachel Neaman, Technology Consultant
- Jeanette Nielson, HR Manager, Belzona Polymers Ltd

- Professor Nic Palmarini, Director, National Innovation Centre for Ageing
- Dr Sukhmeet Panesar, Deputy Director, Strategy and Development (Data and Analytics), NHS England & NHS Improvement
- Reema Patel, Associate Director (Engagement), Ada Lovelace Institute
- Andrew Pearson, Senior Technology Transfer Manager, Innovation Programmes and Intellectual Property, National Institute for Health Research (NIHR)
- Professor Jill Pell, Director, Institute of Health & Wellbeing; University of Glasgow
- Aidan Peppin, Senior Researcher, Ada Lovelace
- Nicola Perrin, Director of Policy and Public Affairs, Association of Medical Research Charities
- Chloe Pont, Data Scientist, Birdie
- Juwairia Quazi, Design Strategist, Planetary Health Lab, University of Edinburgh
- Vincenzo Rampulla, Smart Data, BEIS
- Matthew Reed, Strategy Director, Materials Innovation Factory
- Dr Andrew Roddam, CEO, Our Future Health
- Professor Martin Rossor, NIHR National Director for Dementia Research, Professor of Clinical Neurology, Consultant Neurologist at the National Hospital for Neurology and Neurosurgery, UCL
- Gila Sacks, Head of Prevention, Office for Health Improvement and Disparities
- Damini Satija, Senior Policy Adviser, Centre for Data Ethics and Innovation
- Ash Scott, Director, DoIT.life
- Dr Richard Siow, Director, Ageing Research, King's College London
- Nina Skorytchenko, CEO, Avenna
- Richard Sloggett, Founder & Programme Director, Future Health Research
- Rosie Smith, Senior Government Affairs Manager, Boehringer Ingelheim
- Mr Gavin Starks, CEO & Founder, Dgen
- Adam Steventon, Director of Data Analytics, Health Foundation
- Svitlana Surodina, CEO, Skein
- Joshua Symons, Director of Data Strategy, Genomics England
- Benjamin Szymkow, CEO & Founder, Flusso
- Ming Tang, Chief Data and Analytics Officer, Data and Analytics, NHS England & NHS Improvement
- Christopher Thomas, Centre for Data Ethics & Innovation (CDEI)
- Jim Thomas, Head of Workforce Capacity and Transformation, Skills for Care
- Gerry Toner, COO, Curamcare
- Professor Michael Trenell, Founder & CEO, Changing Health
- Brad Vrabete, CTO, Alwayswithyou Ltd
- Nickolai Vysokov, CEO, BrainPatch.AI
- Mr Peter Ward, Co-Founder & CEO, Humanity
- Michael Warren, Chief Communications Officer, Our Future Health
- Scott Watson, NHS Strategic Digital Advisor, Northern Care Alliance NHS Foundation Trust
- Dr Adrian Weller, Programme Director for Artificial Intelligence, and Turing Fellow, Turing Institute
- Dr Joachim Werr, Founder & Executive Chair, HN
- Professor John Wright, Director, Bradford Institute for Health Research
- Katharine Wright, Assistant Director, Nuffield Council on Bioethics
- Dr Hakim Yadi PhD OBE, Co-Founder & CEO, Closed Loop Medicine Ltd
- Ms Helena Zaum, Social Care Lead (Partnerships), Microsoft UK

## 8. APPENDIX

### 8.1. What are the potential benefits and opportunities?

#### **Benefits**

Improved data sharing informs better decisions, leading to efficiencies in operations as well as growth in user demand for data, products and value-added services (including services relating to the growing data market).

From an economic perspective there is value in moving from Closed data to Shared and Open data, enabling other actors to re-use the data to provide new public and commercial services at the national levels and beyond. In terms of improved efficiency and effectiveness, there are the following benefits:

- Data sharing as a driver of improved data quality;
- Data is more readily discoverable, available, accessible and usable for internal users (as well as external users);
- Reduced overhead (human effort and transaction costs) of managing data exchange through standard licenses rather than multiple, separately negotiated (and hard to enforce) Memorandums of Understanding (MoU);
- Increased availability of data in machine-readable formats increases opportunities for the automation of tasks, releasing skilled staff for other activities which fully utilise their skills – improving both efficiency and staff engagement;
- Improved recognition for the stewardship role undertaken by an institution arising from the implementation of terms and conditions in standard licences for data use which mandate the acknowledgement of data providers;
- Improved data resilience resulting from data sharing; and
- Improved user requirements for processing systems based on a sound understanding of data.

#### **Opportunities**

A number of new/ improved opportunities are likely to result from improved data sharing. Some of these will arise from greater 'ease of working' for an organisation and its stakeholders (including its customers and partners). Meanwhile, the creation of a network of active data users will drive data innovation. Opportunities include:

- Growing demand for data – and growing evidence of that demand and of the value of the data;
- Growing a network of active users of data who can help to iterate the use case, prioritise datasets and maintain the usefulness of the data – prioritising the most widely requested data in the right format and at the highest quality, and enabling 'collective learning and capacity building' for everyone in the network (including staff within implementing institutions);
- Development of new partnerships that can meet the user needs; and
- Development of new services relating to data e.g. training sessions on how to use data.



## 8.2. What are potential risks, blockers and constraints?

There are a number of potential blockers and constraints to the implementation of a model for improved data sharing. These fall under a number of headings and issues include, but are not limited to:

### a) Human

Fear of:

- Data quality not being good enough — fear of embarrassment;
- Losing control of the data;
- Giving data to ‘unauthorised’ users, especially in the case of sensitive data, such as that linked to national security or intellectual property;
- Not being recognised, nor reimbursed, for the role of stewardship and the associated investment of resources (time, skills, finances).

Perception of users’ lack of awareness/ appreciation/ understanding of:

- What the data is, where it comes from and how it is produced (and what other datasets may have been used to produce it); and
- How to use the data ‘properly’ — what can/ cannot be done with the data.

Lack of, or inappropriate, incentives to deliver innovation.

### b) Commercial

- Lack of incentives for regulated utilities to invest, develop/ support open data systems;
- Business cases for sharing specific data, or understanding of its value, are not obvious;
- Desire to maintain competitive advantage relating to the provision of value-added services;
- False expectation that significant revenue can be obtained from selling data which is often not the case (especially for data at low temporal and spatial resolutions), contributing to a reluctance to share data; and
- Investment needs to protect the national interest, security, data rights and/or commercial sensitivities.

### c) Regulatory/ legal

- Uncertain, evolving (cross-sector) regulatory landscape regarding data;
- Compliance concerns and law — including the risk of unnecessarily constraining activities as a result of incorrectly applying a law/ regulation ‘wholesale’; and
- Management of consent, risks and liabilities for sharing data.

### d) Technical

- Lack of resources to supply/ receive data;
- Lack of a secure, controllable environment supporting effective governance;
- Lack of the ability to discover appropriate data according to requirements;
- Lack of access to data in a standardised and structured way;
- Concerns over data quality and a lack of trust in their provenance; and
- Management of structural change and local variation across users.

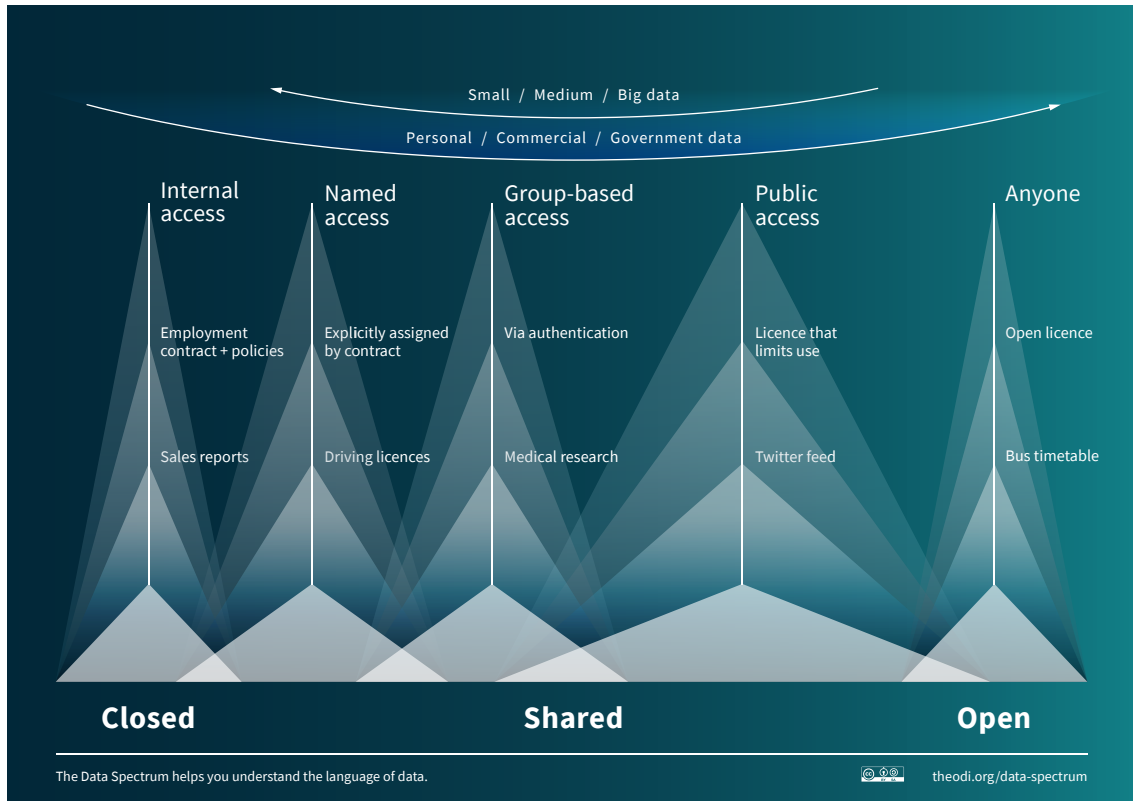
### 8.3. What is data? The Data Spectrum

Open Data can be used by anyone for anything for free (e.g. Creative Commons, Open Government Licence)

Shared Data is data with a preemptive licence (e.g. 'data as a service' that can be used with certain restrictions)

Closed Data requires a user-specific custom licence/contract for use (e.g. 'bilateral contract' for a specific project)

#### The Data Spectrum



50

Data exists on a spectrum of accessibility. The data spectrum ranges from Closed to Shared to Open. The data accessed via an open API may be Closed, Shared or Open data.

<https://icebreakerone.org/open-shared-closed/>

<https://theodi.org/about-the-odi/the-data-spectrum/>

## 8.4. Data maturity models

There are numerous data maturity models.<sup>87</sup> We propose a toolkit that will help unlock organisational capabilities, enabling systemic understanding of user needs and the delivery of Shared Data flows. It is founded on the following guiding principles:

- Build for search to enable discovery;
- Design for open access to discover Shared Data and enable its authorised use;
- Address fair value exchange (reciprocity) in data value chains through licensing; and
- Develop adaptive governance practices to address risk and innovation.

A matrix-management<sup>88</sup> approach to development can provide a simple, but systematic programme to address the five components of action required.

**Strategy:** Defining Key Performance Indicators (KPIs) for reach, influence and impact;

**Processes:** Operationalising publishing, standards development and reuse;

**Skills:** Developing expertise, knowledge and continuous learning;

**Community:** Running engagement process and gathering feedback; and

**Finance:** Budgeting, financial oversight and economic impact assessment.

Data maturity model chart below  
- content at data maturity model 2021-10 for design

Maturity model		Initiated 1	Demonstrated 2	Adopted 3	Operational 4	Learning 5
<b>Strategy</b>	Reach; Influence; Impact	Outcomes planning	Catalogue available Processes validated Skills addressed	Processes adopted Skills embedded	Connections enabled Community engaged	Cost-benefit analysis User-benefit analysis
<b>Processes</b>	Publishing; Standards; Reuse	Planning	Demonstrably re-usable	Used across the organisation	Applied systemically Documented processes	Continuous improvement process
<b>Skills</b>	Expertise, knowledge, learning	Gap analysis	Pioneer staff trained Scoping	Embedded into staff training	Part of professional development	Continuous improvement process fed into training Documented processes
<b>Community</b>	Engagement process; Feedback	Planning Scoping Understanding user needs	Workshops Engagement Training	Whole organisation engaged in learning process	Standardised ecosystem engagement & feedback processes Regular open communications. Standardised codes of practice	Learning feedback loops codified into Processes, Skills and development. Needs-based development. Open sharing
<b>Finance</b>	Budgeting; Oversight; Value Assessments Economic impact Cost-benefit analysis	KPI definition	Procurement rules	Quantitative measure adopted	Delivery against measures	Adaptation to reality Quantification of benefits Investment strategy

87. <https://www.gov.uk/government/organisations/government-data-quality-hub>  
88. [https://en.wikipedia.org/wiki/Matrix\\_management](https://en.wikipedia.org/wiki/Matrix_management)

## 8.5. Licensing frameworks, certification and assurance

There are many existing processes, policies and rules for data licensing. Here we consider public and private goods, free/open and paid/shared models for data sharing. These lists are not exhaustive: they are designed to provoke discussion about how to frame specific licensing conditions against specific use-cases.

User (entities)	Example
Academic	research
Product development	pre-commercial product discovery
Commercial	data-enabled products
Public sector	policy input
Third sector	response, strategic development
Collective action	collective databases
Investors	commercial and non-commercial funding

Restricted use (type)	Variables to limit usage
Jurisdiction	legal area
Temporality	one-off; regular; frequent; real-time
Volume	big; medium; small
Geography	country; region; biome
Dates	allowed for use until; recommended for use before; deprecated after; valid for specific range; licenced between
Application-specific	domain-specific; platform-specific
Project-specific / produced work	scope-limited
Granularity	detail-limited
Aggregated modelling, overlay	use-limited
Compilation vs derivative work	output-limited

For example, predetermined licensing use-cases could include third sector modelling of a specific geography within a limited time frame to address single-issue research. In addition to the use restrictions that could be applied to specific use-cases detailed above, licensing considerations must also be applied to the following areas.

Parameters (type)	Example
Privacy	anonymised; pseudonymised; identified; secret
Confidentiality	level of confidentiality
Security	level of security
Pricing	pricing for a specific scenario or a general case
Ownership	reference to the legal owner
Attribution	legal assignment given
Provenance	provenance modelling (e.g. <a href="https://openprovenance.org/">https://openprovenance.org/</a> )
Authority (unverified)	declared quality (e.g. enterprise, recommended, verified,

One way to consider how to develop minimum viable mechanisms for licensing is to begin with models such as Creative Commons (CC0), for example:

- Public domain (CC0);
- Attribution;
- Share-alike;
- Commercial;
- Non-commercial; and
- No derivatives.

Licences conformant with the Open Definition<sup>89</sup> are listed at [opendefinition.org/licenses](https://opendefinition.org/licenses). Ordnance Survey provides information on common terms used in public sector licences.<sup>90</sup> It is important to note that there are licenses only for Open Data and these are not applicable to Shared Data. Development of licensing conditions for Shared Data is one task that can be addressed via a Trust Framework approach.

### **License representation:**

The EU INSPIRE project recommends that each License be distributed as “Machine readable License”. This term can mean a number of different things:

- It is in electronic form, e.g. a Word document or PDF;
- Adding a tag to a hyperlink to a document to indicate that the document is a licence, e.g. `rel="license"`; and
- That the licence is expressed in xml and/or rdf using an agreed language and template, e.g. ISO/PRF 19149 (GeoREL).<sup>91</sup>
- It is generated automatically by a machine based on input from a potential user. The output of this could be any one of the three possibilities above.

Assurance of data and data practices are important for organisations to build trust, manage risks, and maximise opportunities. Trust frameworks can encourage sharing by helping to de-risk development, opening up opportunities and innovation, reassuring consumers, encouraging better quality control and stimulating behaviour change.

**Data certification** can help build trust in markets. Existing certification systems in the sector can be extended using, for example, existing examples of Open Data Certificates.<sup>92</sup> A potential derivative approach that could provide pointers to users as to a measure of the quality of the data sources they are using.

For example, knowing that a data source will be actively maintained for a number of years can inform a robust investment decision in its utilisation. More broadly, certification can help assess the legal, practical, technical and social aspects of publishing data using best practice guidance. Further, if data descriptions are made available in machine-readable,<sup>93</sup> open formats, certification can be - in part - automated, which can contribute to the stability of the overall data ecosystem.

89. For a more extensive understanding, see: <https://opendefinition.org/od/2.1/en/>  
90. Refer to: <https://www.ordnancesurvey.co.uk/business-and-government/help-and-support/public-sector/guidance/glossary.html>

91. See: <http://www.iso.org/iso/rss.xml?csnumber=32567>

92. More information at: <http://certificates.theodi.org>

93. A primer on machine-readable data: <https://www.data.gov/developers/blog/primer-machine-readability-online-documents-and-data>



### Categories related to data publishing & certification

Release types	Can individuals be identified?	Supported
one-off	no, not about individuals	human contacts
ongoing	no, anonymised	discussion mechanisms (public and private)
serviced (continuous)	yes, or risk of	tools that support usage

Shelf life	Type	License
will expire at some point closed;	primary source data	type (open, shared, derivatives)
has expiry date	derivative	context (use-case)
will not expire	part of a collection	context (organisation)

Availability guarantee	Formats	Guarantees
not defined	machine-readable	backed up
6 months	standard open format	distributed/ multihomed
a year	human-readable	
indefinitely	persistent identifiers	

If structured	Quality	Rights
semantic (HTML, Markdown, Docbook)	no process	rights assertion
information on formatting (OOXML, PDF)	data documented (e.g. schemas)	links to licenses (e.g. derivatives)
statistical (SDMX, Data cube)	provenance	
structured (CSV, JSON, XML)	quality process documented	
non-structured (Word, PDF)	peer-reviewed	



# The Health Foundation

**This project was supported by the Health Foundation.  
The Health Foundation is an independent charity committed to bringing about better  
health and health care for people in the UK.**



All-Party Parliamentary Group  
Longevity

# Open Life Data Framework

