



Improving the use and utility of UK routine data

UK-Japan symposium on data-driven health 26 February 2020, London



25 February 2020



国立研究開発法人 日本医療研究開発機構
Japan Agency for Medical Research and Development

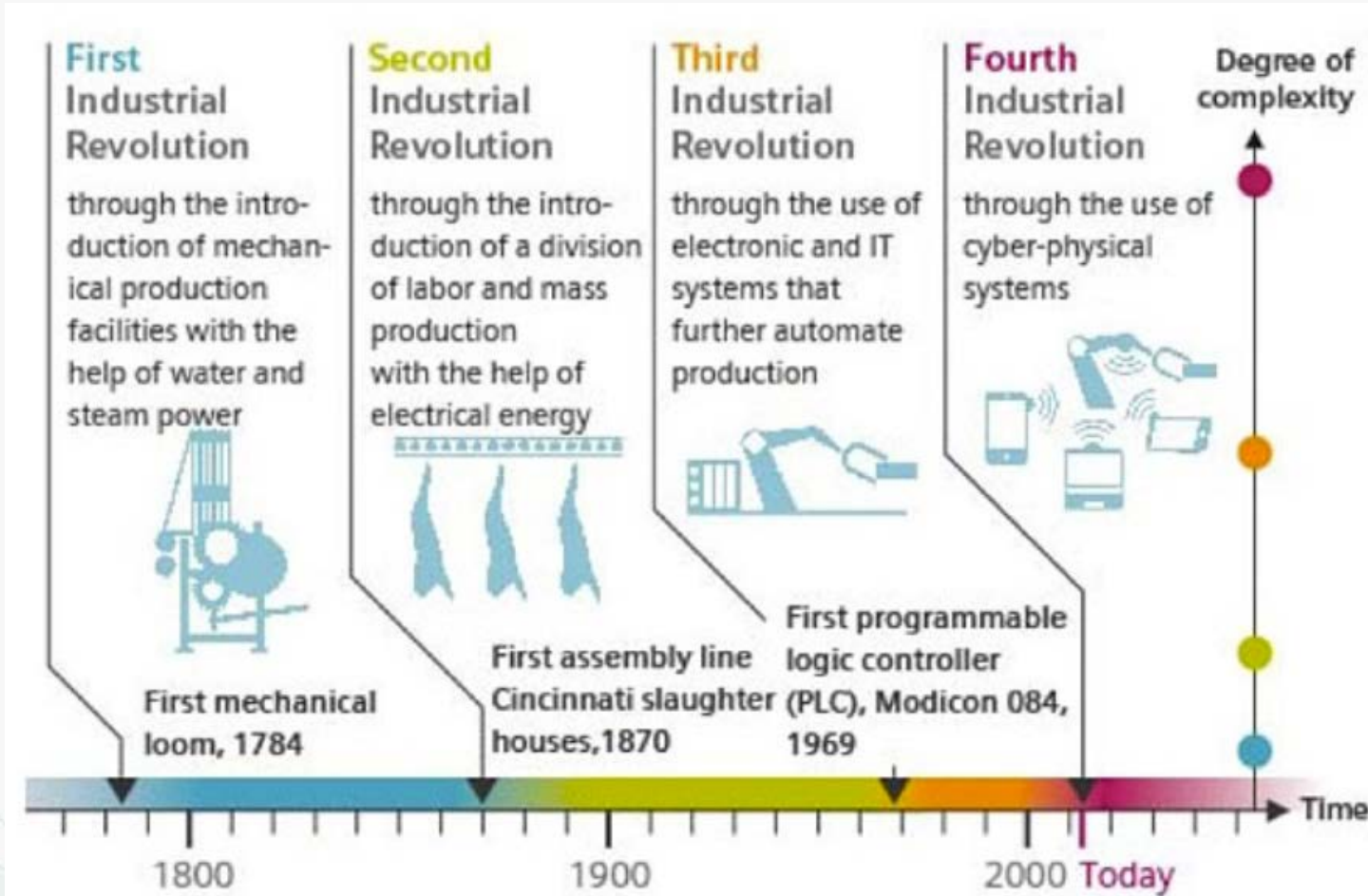


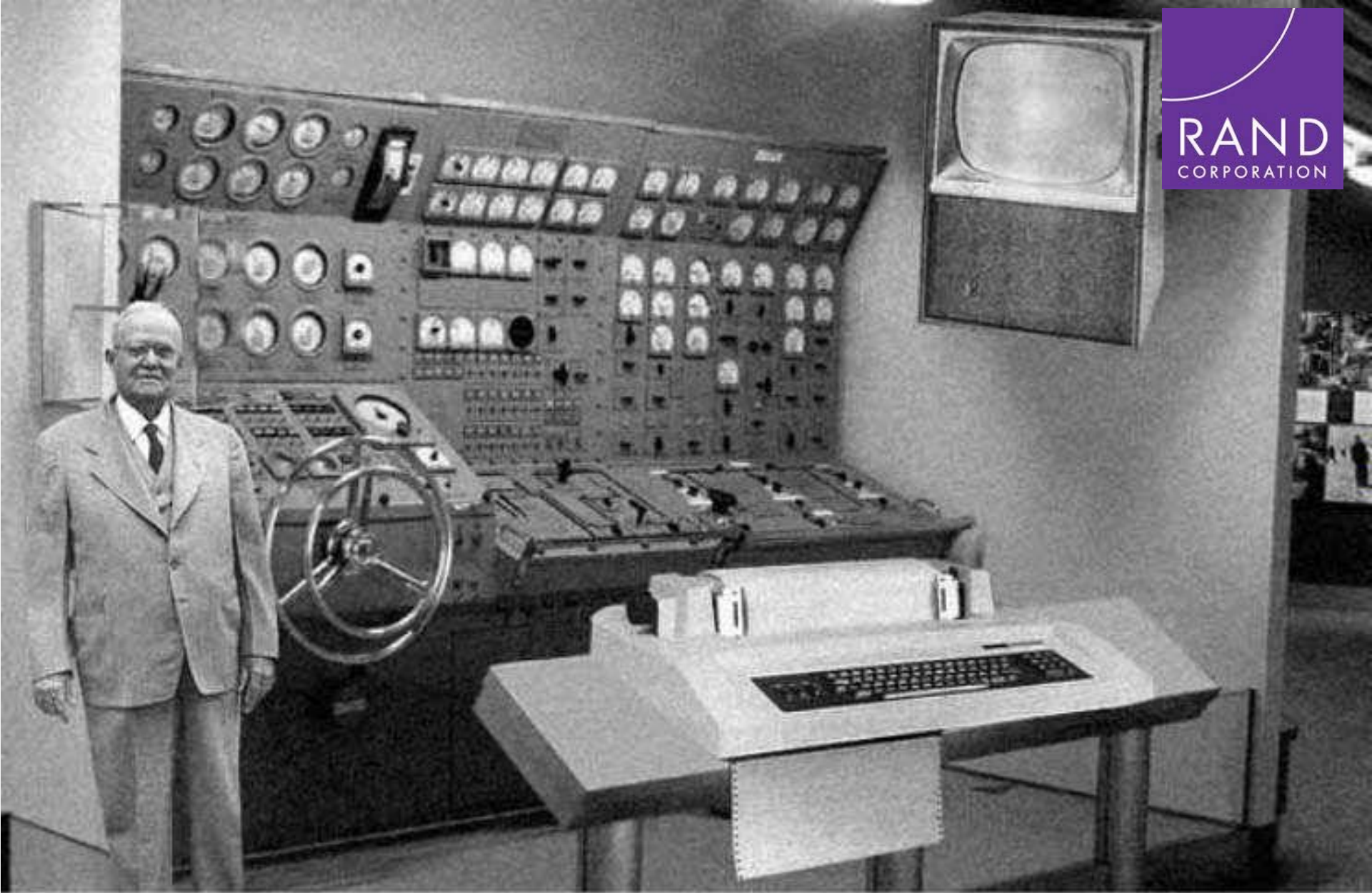
The Next 15 minutes

- **Gearing an entire country for quality health care, research and innovation**
- **Data Science as a Catalyst for Change**
- **With big data goes big responsibilities**



The 4th Industrial Revolution



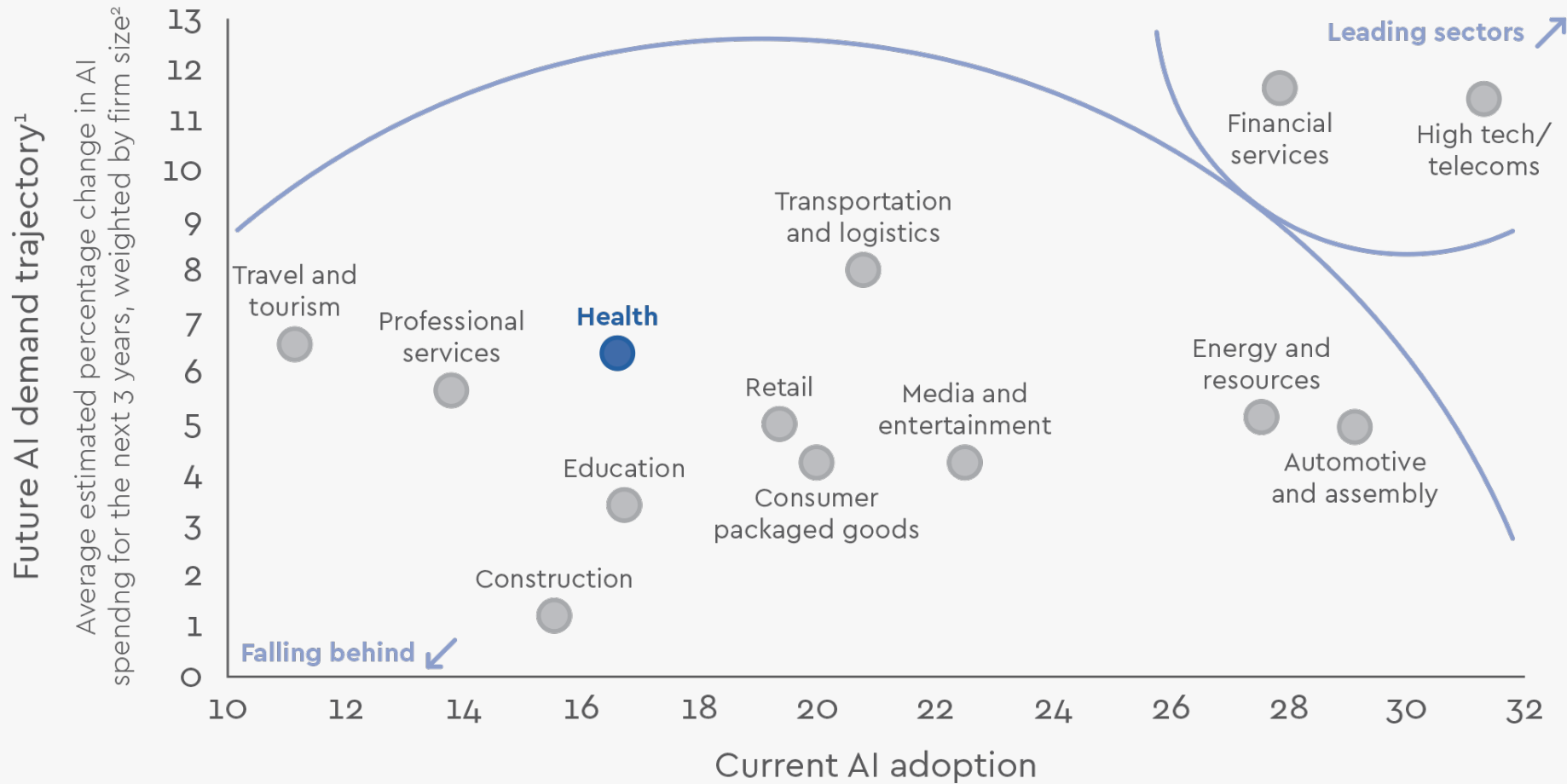


RAND
CORPORATION

DRUK
Data Research UK

Scientists from the RAND Corporation have created this model to illustrate how a "home computer" could look like in the year 2004. However the needed technology will not be economically feasible for the average home. Also the scientists readily admit that the computer will require not yet invented technology to actually work, but 50 years from now scientific progress is expected to solve these problems. With teletype interface and the Fortran language, the computer will be easy to use.

Sectors leading in Data Science and AI adoption today



Percentage of firms adopting one or more AI technologies at scale or in a core part of their business, weighted by firm size²

SOURCE: McKinsey Global Institute, AI adoption and use survey.
 1 Based on the midpoint of the range selected by the survey respondent
 2 Results are weighted by company size.



HDR UK's mission is to unite the UK's health data to enable discoveries that improve people's lives

Our 20-year vision is for large scale data and advanced analytics to benefit every patient interaction, clinical trial, biomedical discovery and enhance public health.

About Us

Operating Model

- Ten Funders - **we are not a funder**
- Separate legal entity, tiered structure,
- Wellcome Trust base – **not a data controller**
- “Lean”

Substantive Scientific Sites

- Science, leadership, NHS partnership and delivery

Single set of Terms and Conditions for collaborations

- Core platforms, inter-operability, standards, governance, meta-data dictionaries
- Partnership – coordinating and associate ROs
- £37M now £120M Initial Investment

Early days!

- - 520 days old; UK Biobank 4894 days old

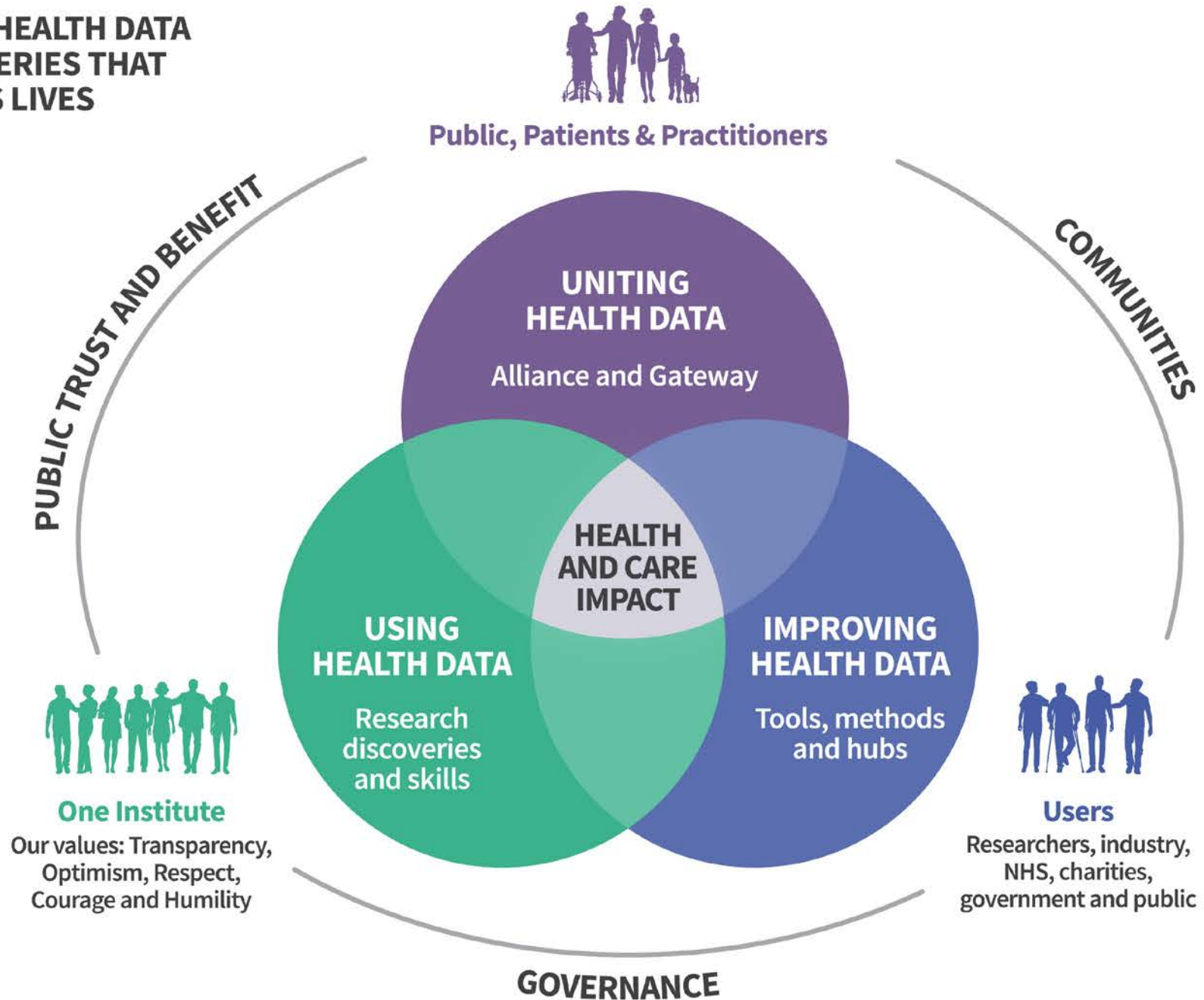


Partnerships at scale

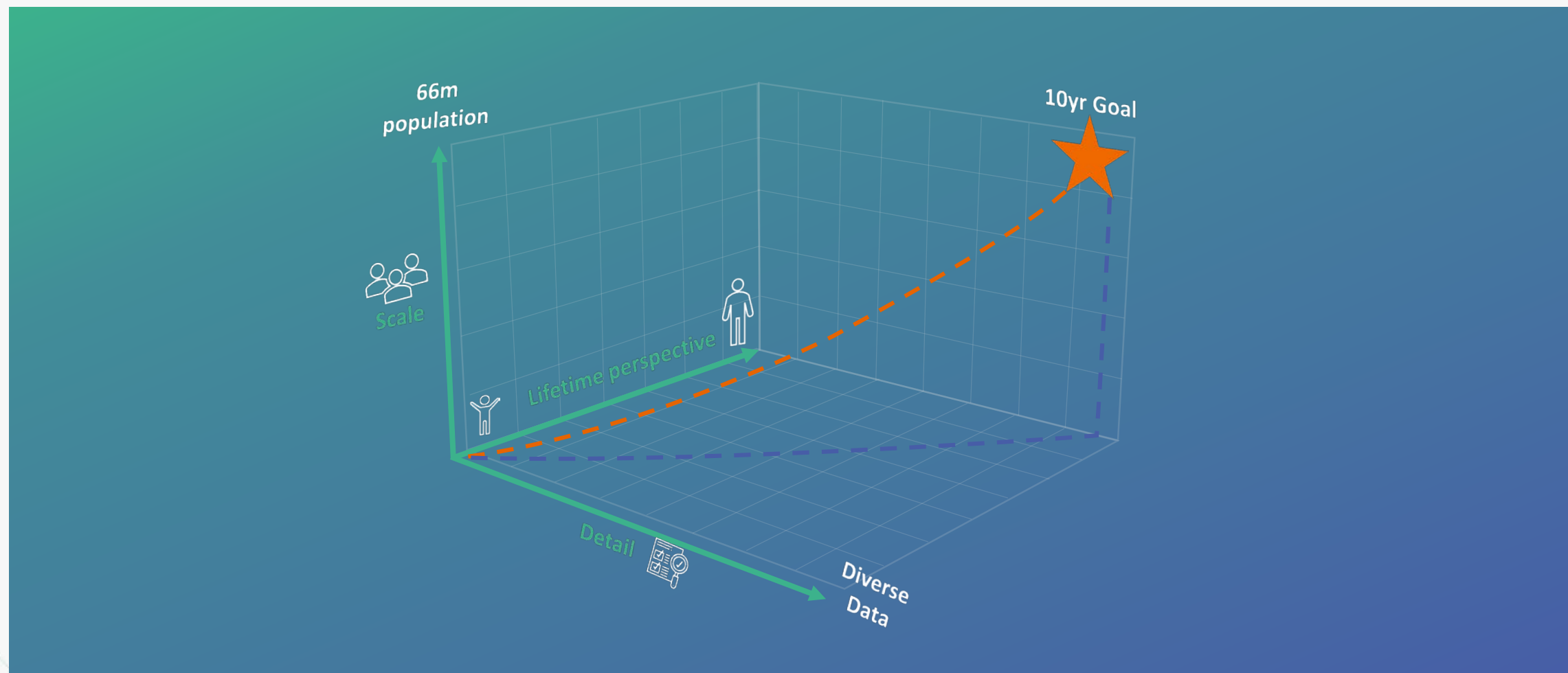
Our initial investment supports six research sites. Each has world-class expertise; a track record in using health data to derive new knowledge, scientific discovery and insight; and works in close partnership with NHS bodies, industry and the public to translate research findings into benefits for patients and populations.

	1. Wales and Northern Ireland (Swansea and Queen’s University Belfast)
	2. Midlands (Birmingham, Leicester, Nottingham, Warwick)
	3. Scotland (Glasgow, Edinburgh, Dundee, Aberdeen, Strathclyde, St Andrews)
	4. London (Imperial, Kings, London School of Hygiene and Tropical Medicine, Queen Mary, UCL)
	5. Oxford
	6. Cambridge (EBI, Sanger, Cambridge University)

**UNITING THE UK'S HEALTH DATA
TO ENABLE DISCOVERIES THAT
IMPROVE PEOPLE'S LIVES**



Our 10-year ambition for scale



Improving the use and utility of UK routine data

A Robust UK Health Data Research Infrastructure



INDUSTRIAL
STRATEGY

UK Research
and Innovation

HDRUK
Health Data Research UK

- Industrial Strategy for Life Sciences (November 2017) Hubs for health data research (*digital innovation hubs*)
- Funded as part of the Industrial Strategy Challenge Fund (ISCF) *Data to Early Diagnosis and Precision Medicine* Challenge
- Health Data Research UK asked to lead delivery of this programme on behalf of UKRI in September 2018
- Ensure that the enormous potential value of the UK's health data assets can be realised
- Four-year programme



Working in partnership with patients, the NHS, universities, business & charities to create a world-leading and robust health data infrastructure



Gearing up the UK for quality health care, research and innovation



Enabling data science and innovation as a catalyst for change



But with big data comes with big responsibilities!

We've been listening and designing a model that's fit for the future of health data research



We've engaged with over
2,700
people

We've spoken to over
350 organisations
in the UK and globally
across NHS, industry
and academia

35

Events

32

Interviews
with Businesses



Inspired by global collaborations

No single organisation possesses all the information required to address every health and well-being issue facing society.

Timely access to data across organisations requires:

- Expertise
- Trusted governance
- Interoperability



Inspired by...



Global Alliance
for Genomics & Health
Collaborate. Innovate. Accelerate.



W3C[®]



“As a neutral global cooperative, SWIFT is defined by its community of users around the world”.



We are delivering what patients expect and what industry, NHS and academic data users want



Data: Longitudinal, event-based, multi-modal, curated datasets on disease sub-groups



Speed: Fast response times and streamlined contracting, governance and approvals



Access: Single, easy-to-use route to data



Expertise: Access to domain specialists, AI and applied analytics, phenotyping



Scale: High quality data with UK-wide coverage



Multiple purposes: Real world data, ability to rapidly identify trial cohorts

All underpinned by building people's confidence and trust in how data is used

A strong public advisory board

Patient, carer and public voice embedded in governance via Public Advisory Board



Kush Kanodia



Rachel Plachcinski



Margaret Rogers



Ben Johnson



Sarah Brooke



Claire Cooper

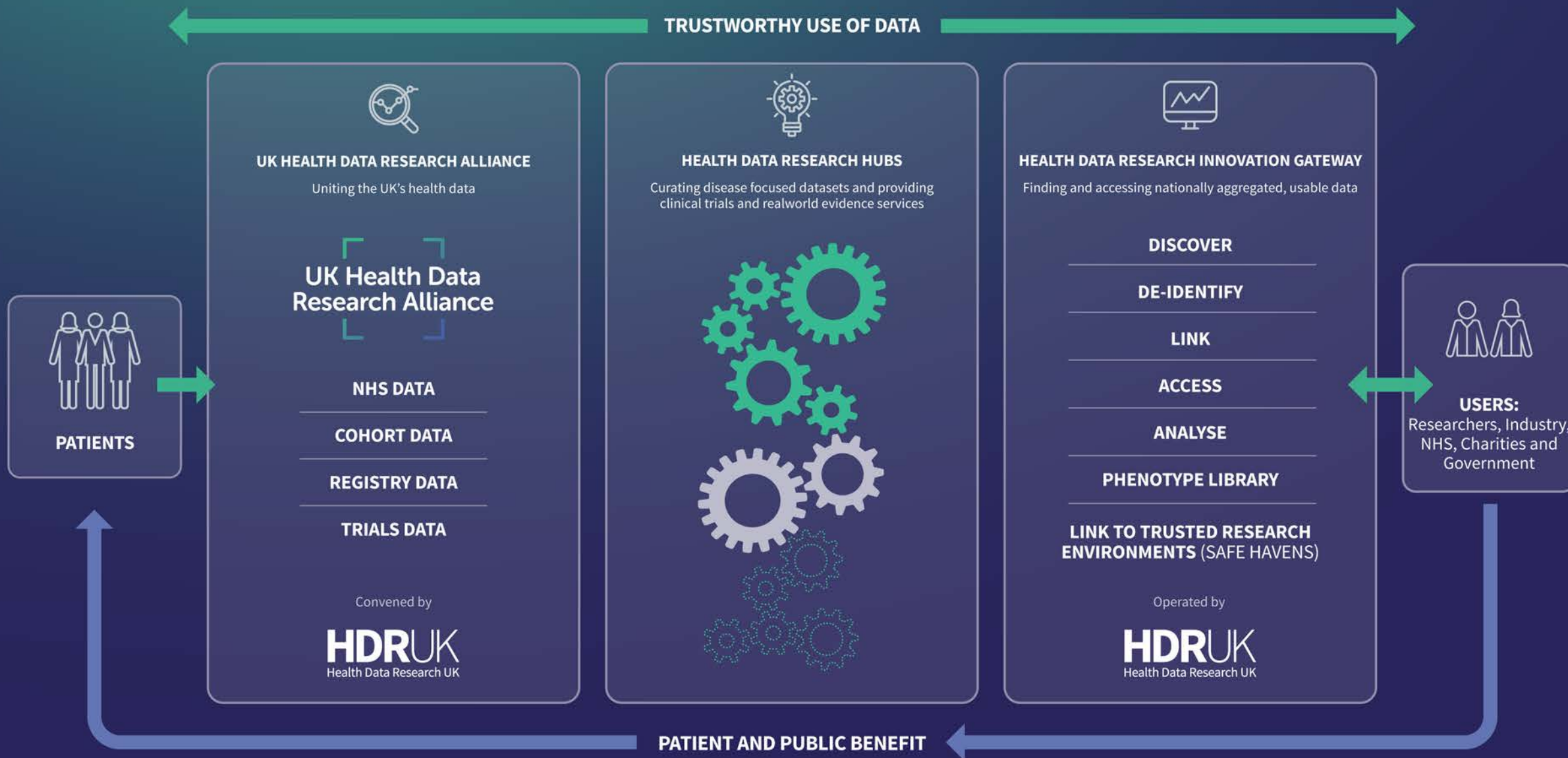


Angela Coulter



Colin Wilkinson

UNITING THE UK'S HEALTH DATA TO MAKE DISCOVERIES THAT IMPROVE PEOPLE'S LIVES



What we have been doing since September 2018



Four-year programme launched in September 2018



UK Health Data Research Alliance launched in February 2019



DIH prospectus launched May 2019

7 Hubs launched September 2019 and started 1 October 2019



Milestone 1 met by 100% hubs 31 December 2019



Livestreamed event: Realising patient and NHS benefits from health and care data:

From policy to practice 22 January 2020



Gateway MVP launch with >400 datasets February 2020

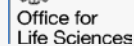
Gateway Rapid Development Task January – March 2020

September 2018

2019

2020

Design & dialogue September 2018 – April 2019



11 Sprints March 2019 – Dec 2019



Industry engagement events UK, US, Europe October 2019 onwards



27 Alliance members announced January 2020



Alliance Symposium 4 February 2020



Alliance



Strategic Partnerships – NHS Digital



- Noel Gordon, Sarah Wilkinson, Tom Denwood

- Priority areas:

- Joint strategy on research data services
- Infrastructure
- Training
- Trials
- Governance



An alliance of leading healthcare and research organisations united to establish best practice for the ethical use of UK health data for research at scale.

Knotty Issues

- Information governance data access*
- Data quality*
- Sustainability and commercial models*
- Meaningful public engagement*
- Technology strategy and standards*
- Inter-operability*

UK Health Data Research Alliance in numbers

27 Alliance members

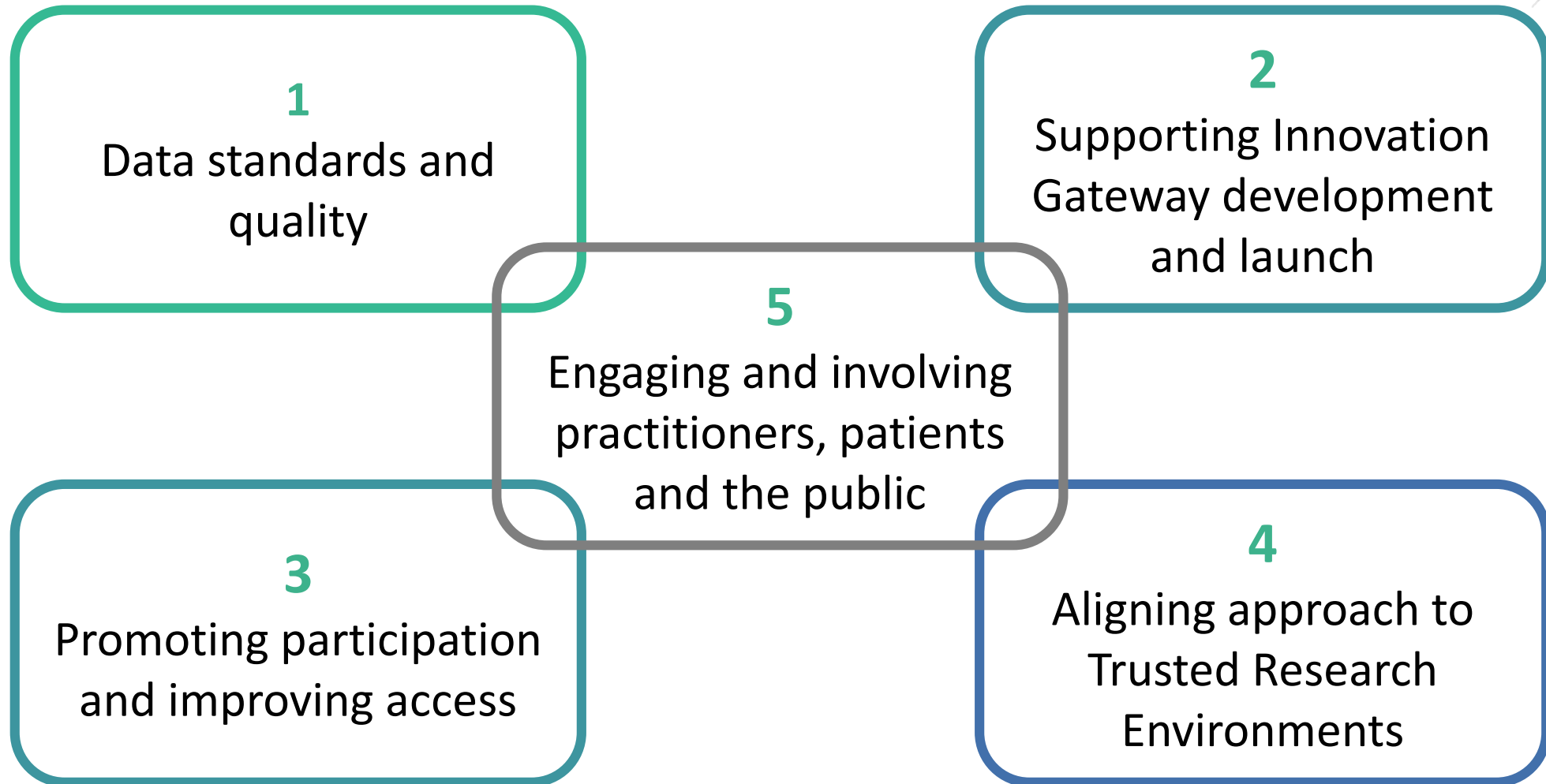


5 working sub-groups

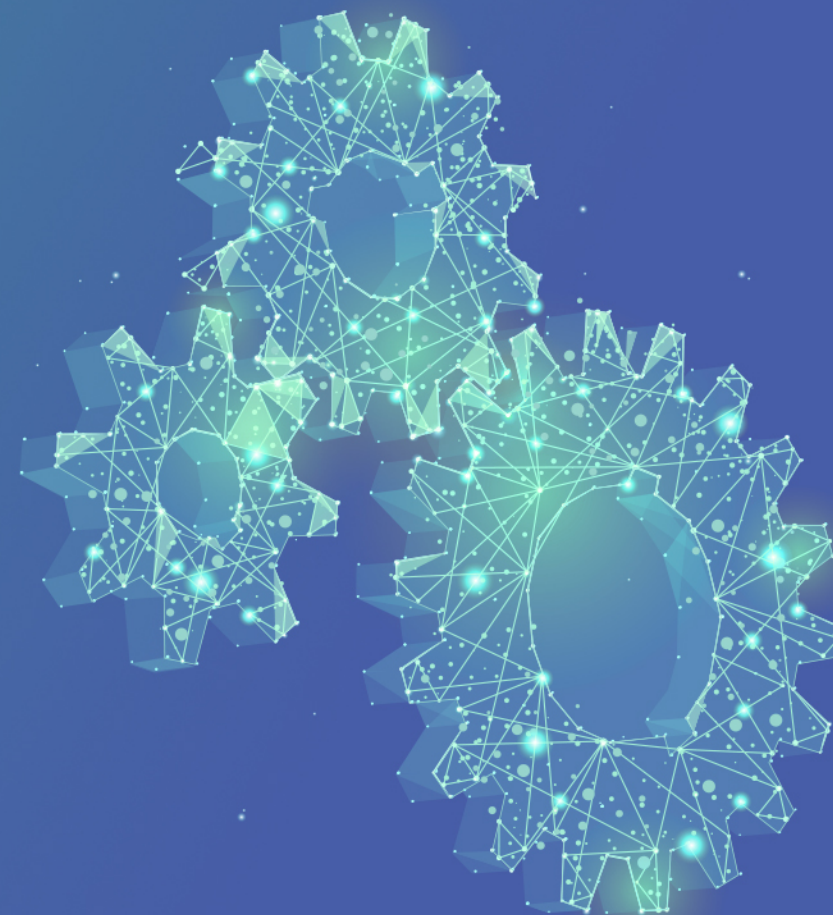
- **Data Quality & Standards** convening **>30** data officers
- **Practitioner, Public & Patient Involvement and Engagement (PPPIE)**
- Supporting development of the **Innovation Gateway** with **>400** datasets discoverable
- **Promoting participation and improving access** **>300** viewers of live streamed event
- **Aligning approach to Trusted Research Environments**



UK Health Data Research Alliance workstreams



Hubs



Applications will be considered by the panel based on pre-determined criteria

- **Make datasets available** for research, development and innovation across industry, academia and the NHS
- **Curate new or existing data** for research, development and innovation taking place across industry, academia and the NHS
- **Provide services** to meet identified needs for research, development and innovation taking place across industry, academia and the NHS
- **Deliver** against agreed milestones
- Engage with and involve **patients and the wider public**
- Demonstrate **capabilities and partnership**
- Enable **wider impacts** and linkage to the aims of ISCF and the Life Sciences Industry Strategy
- Provide **value for money**

7 Health Data Research Hubs



UK Health Data Research Hubs in numbers

100% of the **7** hubs met milestone 1;
involving **>100** organisations from NHS,
academia, industry and charities involved
in **22** locations across England, NI, Scotland &
Wales

>4 high profile global partnerships

Discover-NOW
AstraZeneca

INSIGHT
Health Data Research Hub
Roche

Gut Reaction
Health Data Research Hub
Microsoft

NHS DigiTrial
NOVARTIS

6 disease areas prioritised

Cancer
Diabetes
Respiratory health
Inflammatory bowel disease
Eye health
Acute care

Quality of datasets

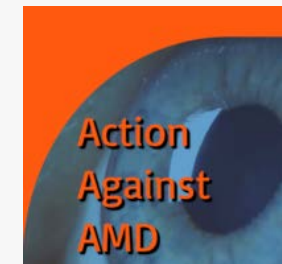
★★★★☆ **Level 4: Full technical submission**

Demonstrated technical maturity and capability to submit
quality metadata

- DATA-CAN
- DiscoverNOW
- INSIGHT
- Gut Reaction

Charity Partners in Hubs

Charities are mobilising the community to share, celebrate and challenge the work of the Hubs; directly supporting research by providing access to data asset; bringing in patient voices



• 12 September, BMA House

Health Data Research Hubs Announcement



a. Partnerships spanning across the UK

INSIGHT -

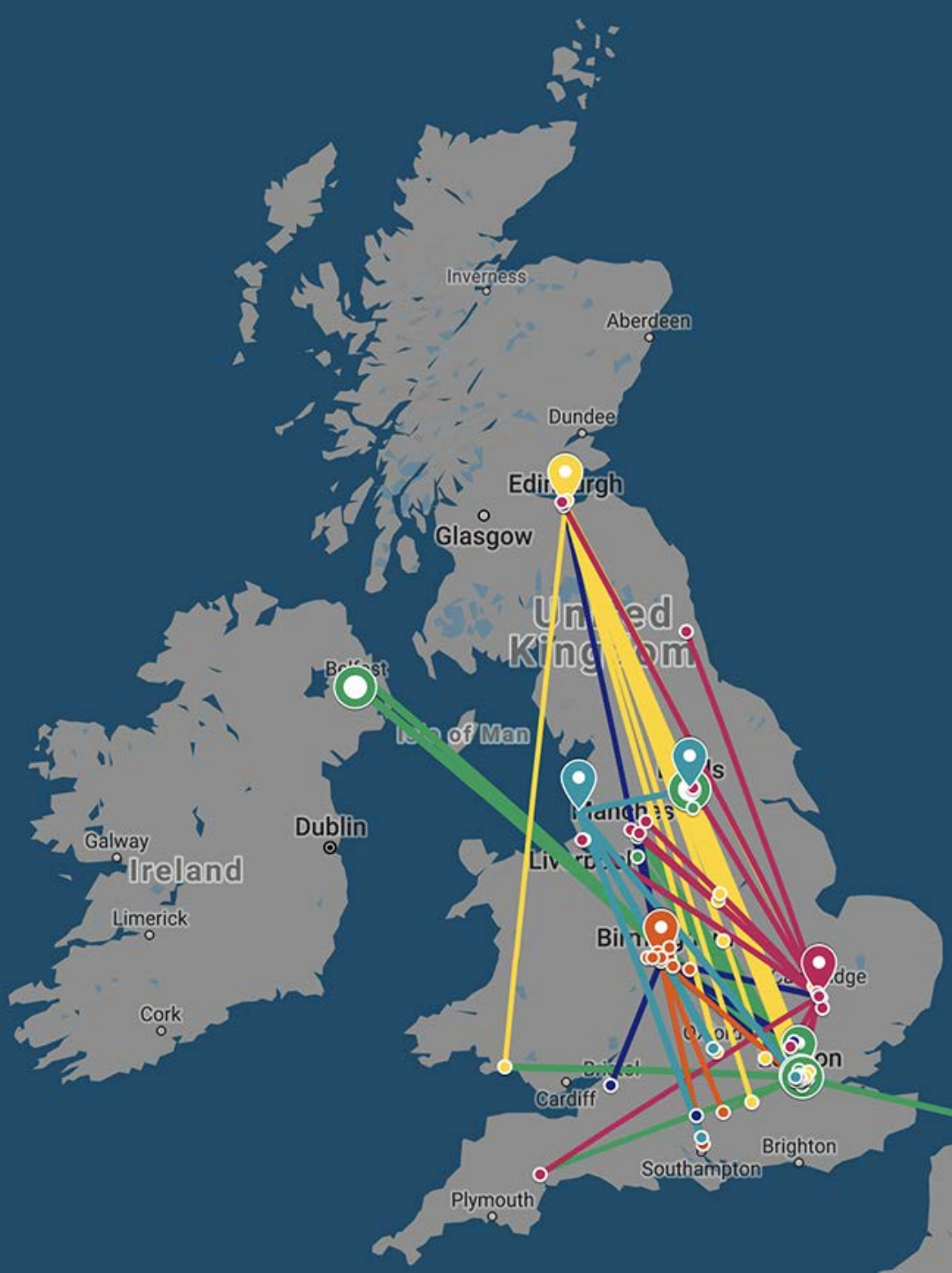
<https://www.youtube.com/watch?v=BkPXFfetZwM>

PIONEER -

<https://www.youtube.com/watch?v=eibnqYye0og>

NHS DigiTrial -

<https://www.youtube.com/watch?v=dRGrUusoqVE>



Example: The Discover Now Hub builds on one of Europe's largest and most diverse, longitudinal linked data sets

2.3
million
population



More than
10
years of
historic data



>£20
million
invested



6
Terabytes
of data



Over
400
ISAs signed
with partner
organisations



Our partners

Technology Partners



Industry Partners



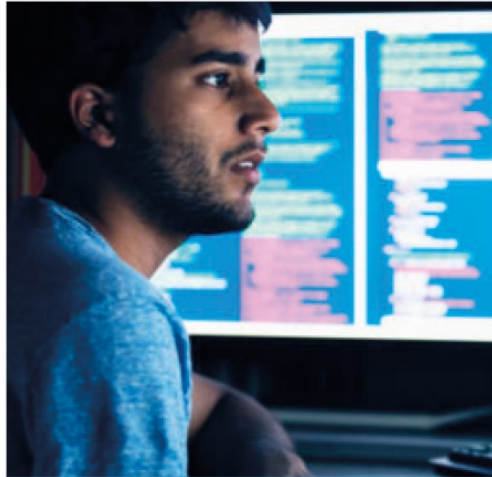
NHS, Academic and Third Sector Partners



Data and Research Partners



Example: Discover Now Hub - Four domains to create a world leading RWE Hub



Richer and curated data

- **Additional data** including user generated
- Growing across London from **2.3 to 8 million** lives
- Improved quality of data



Leading technology

- Moving to **near real time** and high performance analytics
- **Single-point, streamlined access** to de-identified real-world data for trials



Implementation ready

- Embedded in a **real-world care eco-system**
- Access to leading KOLs and data and research expertise



Public trust

- Ambitious London wide **deliberative democracy** approach to build trust
- Commitment to be guided by citizens recommendations

Our partners

Technology Partners



Industry Partners



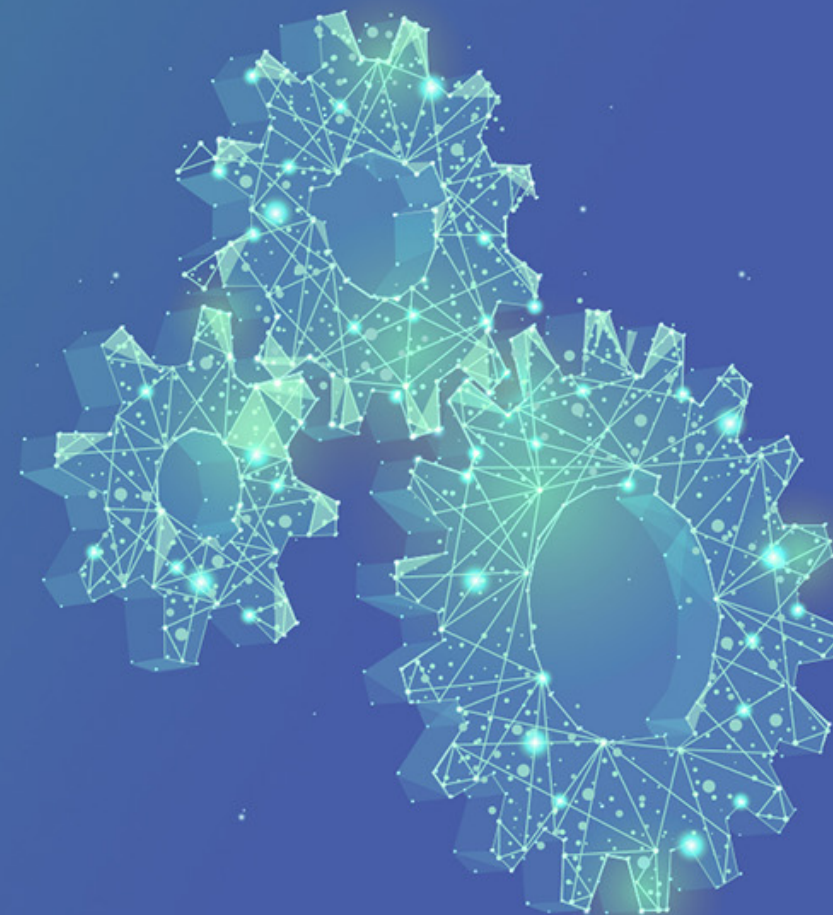
NHS, Academic and Third Sector Partners



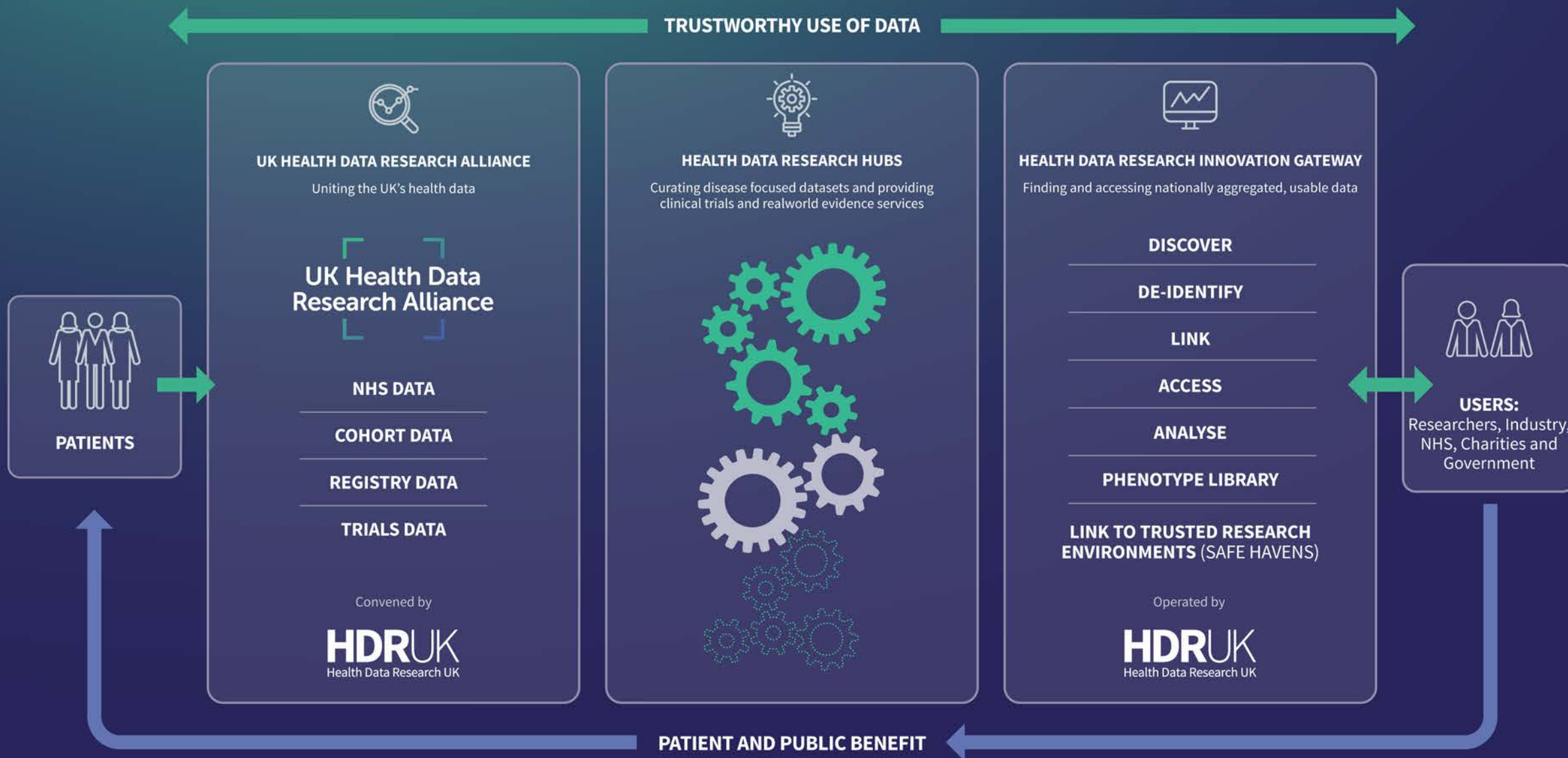
Data and Research Partners



Innovation Gateway

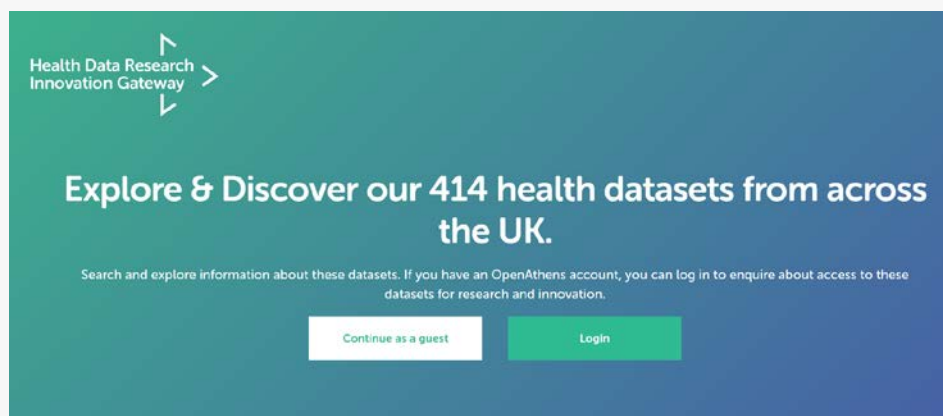


UNITING THE UK'S HEALTH DATA TO MAKE DISCOVERIES THAT IMPROVE PEOPLE'S LIVES

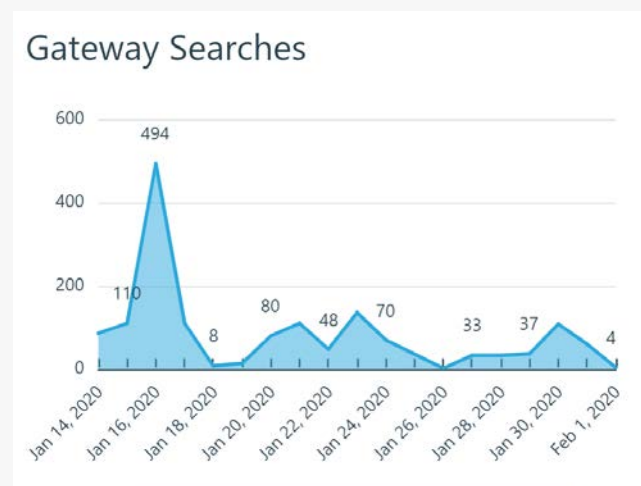


UK Health Data Research Innovation Gateway in numbers

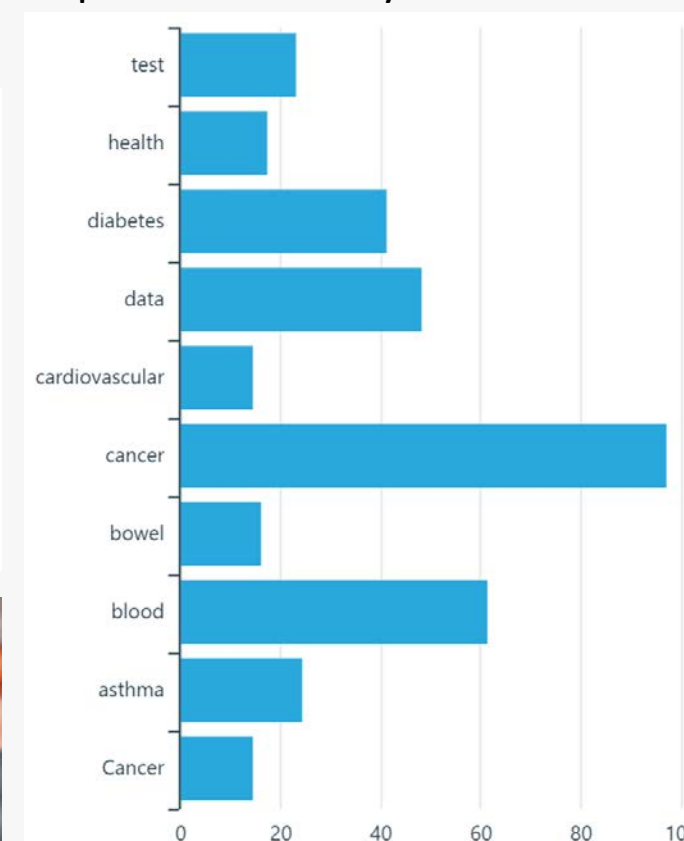
>400 datasets in **Minimum Viable Product**



2,415 searches



Top **10** Gateway searches



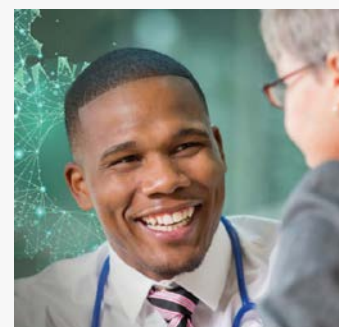
16 organisations registered interest for the **Technology Partnership**

7 applied

3 selected for Rapid Development Task

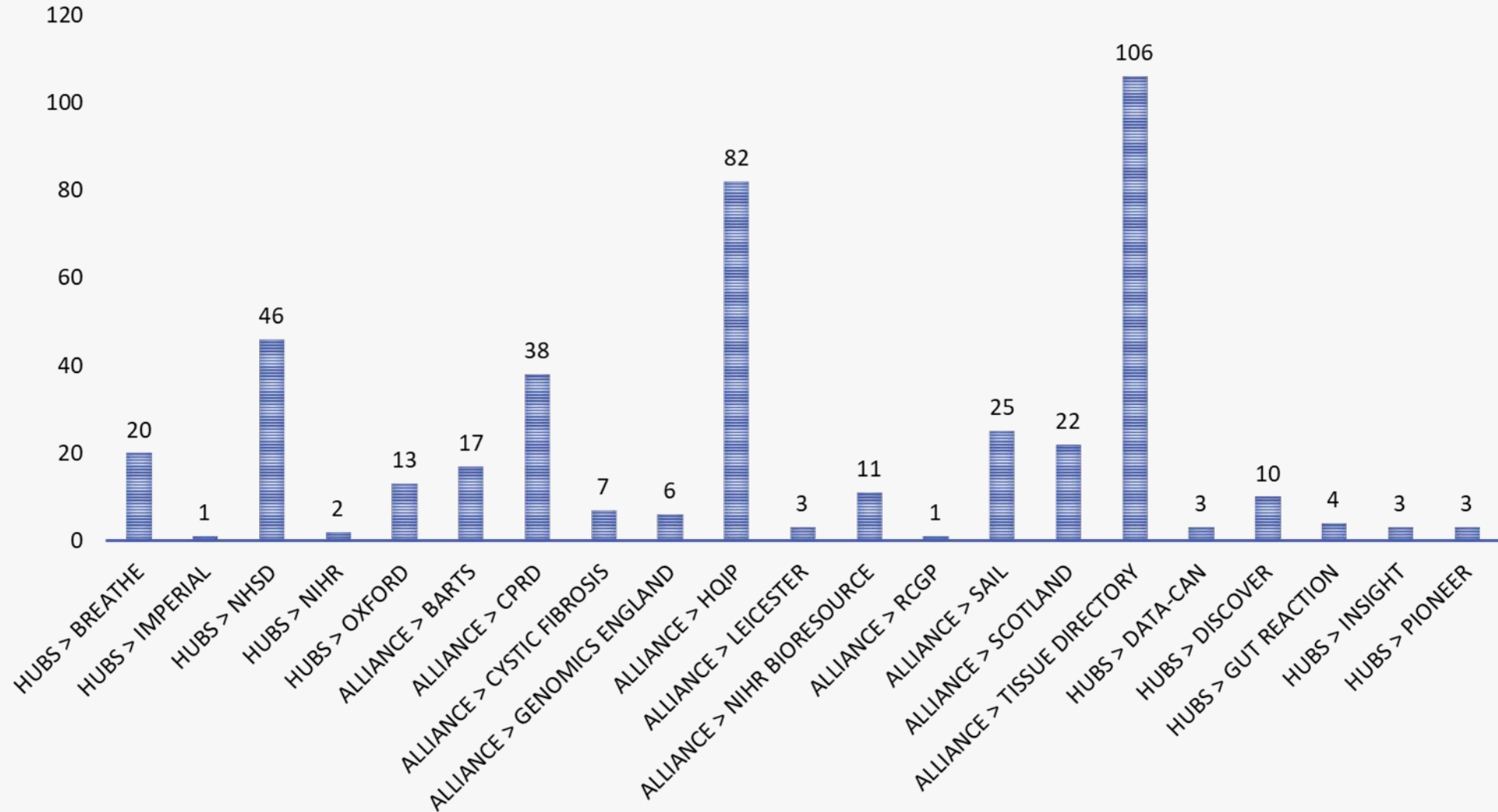
2,076

users

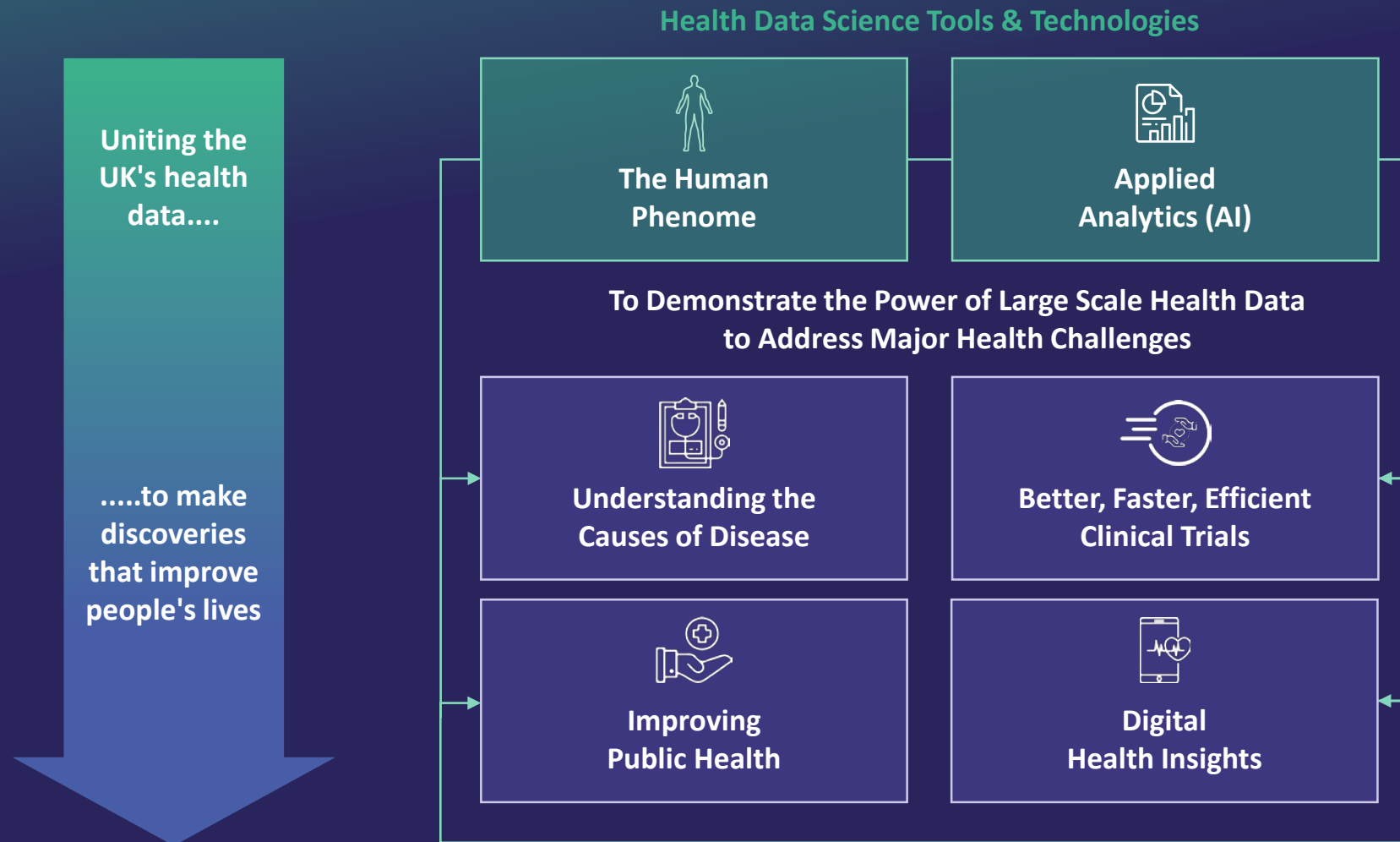


Where the data has come from

Number of datasets on Gateway, February 2020



Scientific Exemplars



Scientific Priorities for 2019/20 and Beyond

Health Data Science Tools & Technologies

Cathie Sudlow
HDR UK Scotland

Harry Hemingway
HDR UK London




The Human Phenome



Applied Analytics (AI)



Chris Holmes
Alan Turing Institute

To Demonstrate the Power of Large Scale Health Data to Address Major Health Challenges

John Danesh
HDR UK Cambridge




Understanding the Causes of Disease



Better, Faster, Efficient Clinical Trials



Martin Landray
HDR UK Oxford

Ronan Lyons
HDR UK Wales/NI




Improving Public Health



Digital Health Insights



Simon Ball
HDR UK Midlands

**The
Alan Turing
Institute**



Applied Analytics

Chris Holmes





APPLIED ANALYTICS (AI)

HDR UK National Text Analytics Resource



Leads: Angus Roberts
Richard Dobson
KCL
Start: 1st Oct 2019
End: 31st Mar 2023
Award: £0.65m (3.5yrs)



80% of the world's data is unstructured

National involvement:



Scotland
Cathie Sudlow (UoE, Neurology/ Epidemiology) ●●
Honghan Wu (UoE, HDR UK Fellow) ●●
Bea Alex (UoE, NLP) ●

Midlands
Simon Ball (UoB, Digital healthcare) ●
Georgios Gkoutos (UoB, Informatics) ●●

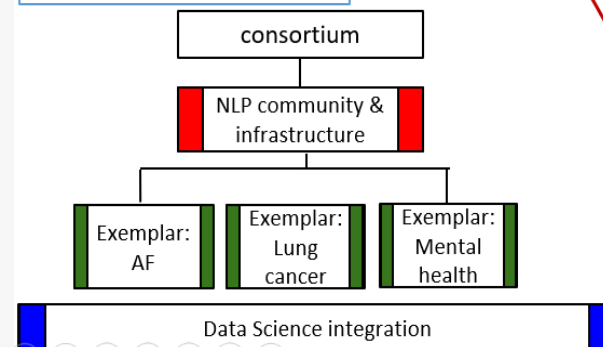
Cambridge
Nigel Collier (NLP) ●

Wales and NI
Ronan Lyons ●●
(Swansea, Epidemiology)
Simon Thompson ●●
(Swansea, Data Science)

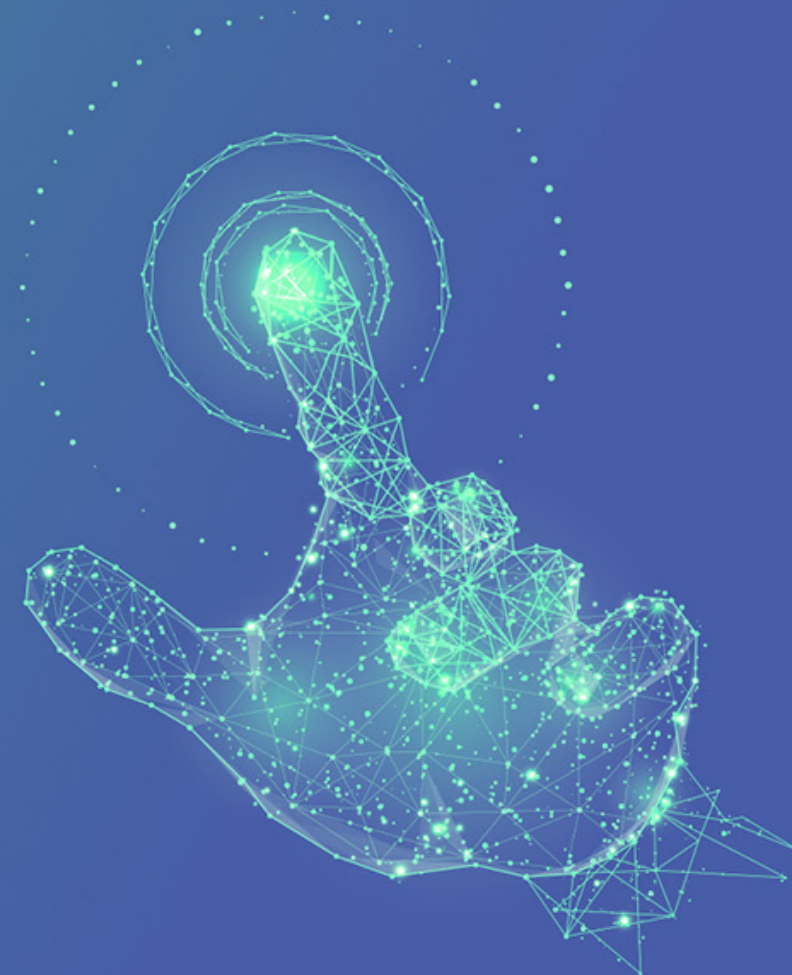
London
Richard Dobson (KCL/UCL, Health Info) ●●
Angus Roberts (KCL, NLP) ●
Robert Stewart (KCL, Psychiatry) ●
Carol Dezateux (QMUL, Epidemiology) ●

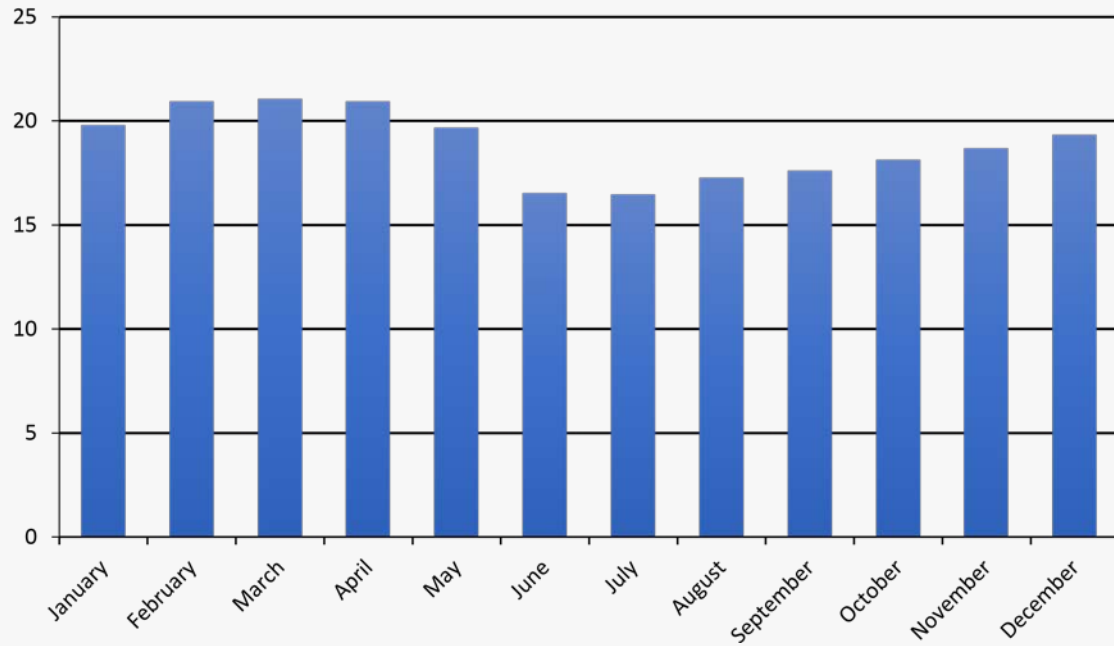
Main aims:

- 1) Survey and map clinical Natural Language Processing (NLP) landscape in the context of EHR analysis – in doing so, build NLP community
- 2) Build a portal for NLP tools, datasets, APIs and standards
- 3) Develop and test tools in stroke, lung cancer and serious mental illness



Improving Public Health





Crude prevalence of learning disabilities by month of conception

UVB quintile	OR	95% CI	P value
1 Lowest	1.00		
2	0.84	0.80-0.88	<0.01
3	0.61	0.57-0.66	<0.01
4	0.55	0.51-0.60	<0.01
5 Highest	0.50	0.46-0.55	<0.01

Multivariate model of association between whole pregnancy UVB exposure and learning disabilities

Conceiving in winter raises learning disability risk

Pregnant women who don't get enough sunshine 'are more likely to have children with dyslexia, autism or ADHD', major study suggests

- 21% of children conceived in February, March and April had a learning disability
- In contrast, the figure was just 16.5% for youngsters conceived in June and July
- The findings come from more than 400,000 youngsters living in Scotland
- Experts behind the study say a lack of sunlight in pregnancy could be to blame

Cross-sectoral data linkage:

- Education (school pupil census)
- Health (maternity records)
- Environment (NASA satellites - solar radiation)



SCIENTIFIC REPORTS

OPEN Antenatal exposure to solar radiation and learning disabilities: Population cohort study of 422,512 children

Received: 12 April 2019
Accepted: 7 June 2019
Published online: 27 June 2019

Claire E. Hastie¹, Daniel F. Mackay¹, Tom L. Clemens², Mark P. C. Cherrie², Albert King³, Chris Dibben⁴ & Jill P. Pell¹





IMPROVING PUBLIC HEALTH

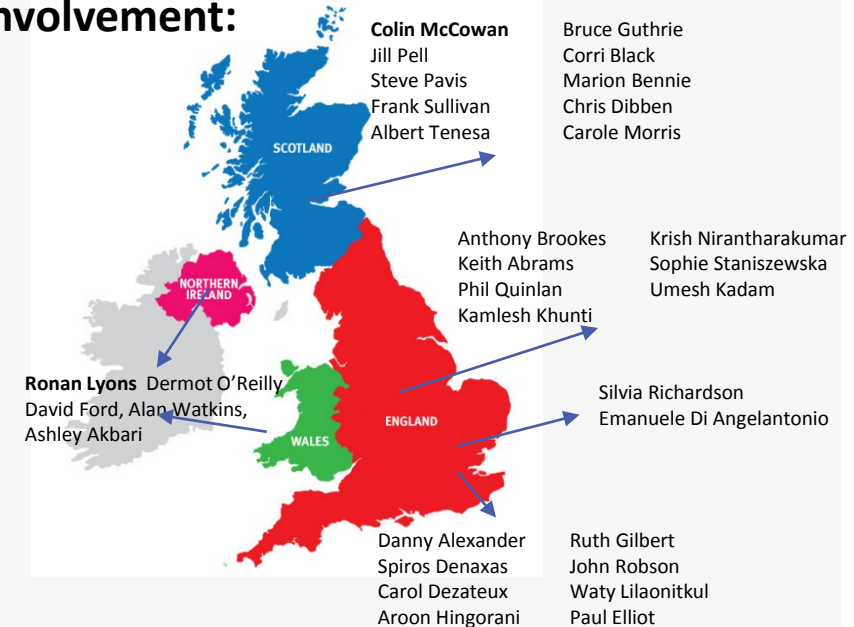
HDR UK National Multimorbidity Resource



Data linkage:

Data Provider	Type	Base Population
Scotland: eDRIS	Hospital Prescribing Disease registries Population spine	5.4m
Northern Ireland	Hospital Prescribing Disease registries Population spine	1.2m
Wales: SAIL	Hospital Prescribing Disease registries Population spine GP	3.5m
CALIBER	Hospital Disease registries GP	10m
Leicester City, Leicestershire & Rutland	GP	1.1m
Discovery (London)	GP Hospital	1.2m

National involvement:



Lead: Colin McCowan
St Andrews

Start: 1st Oct 2019

End: 31st Mar 2021

Award: c£1m (1.5yrs)

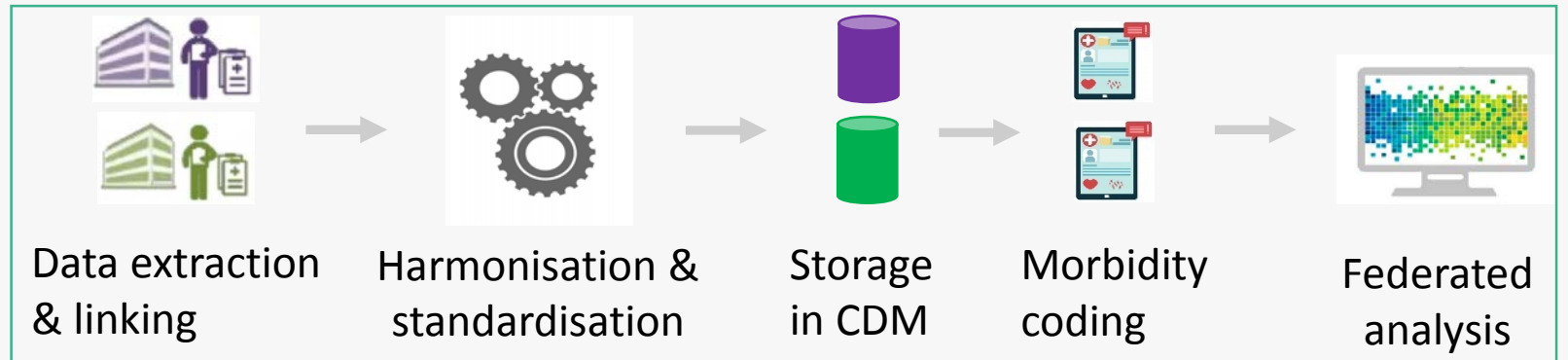


Main aims:

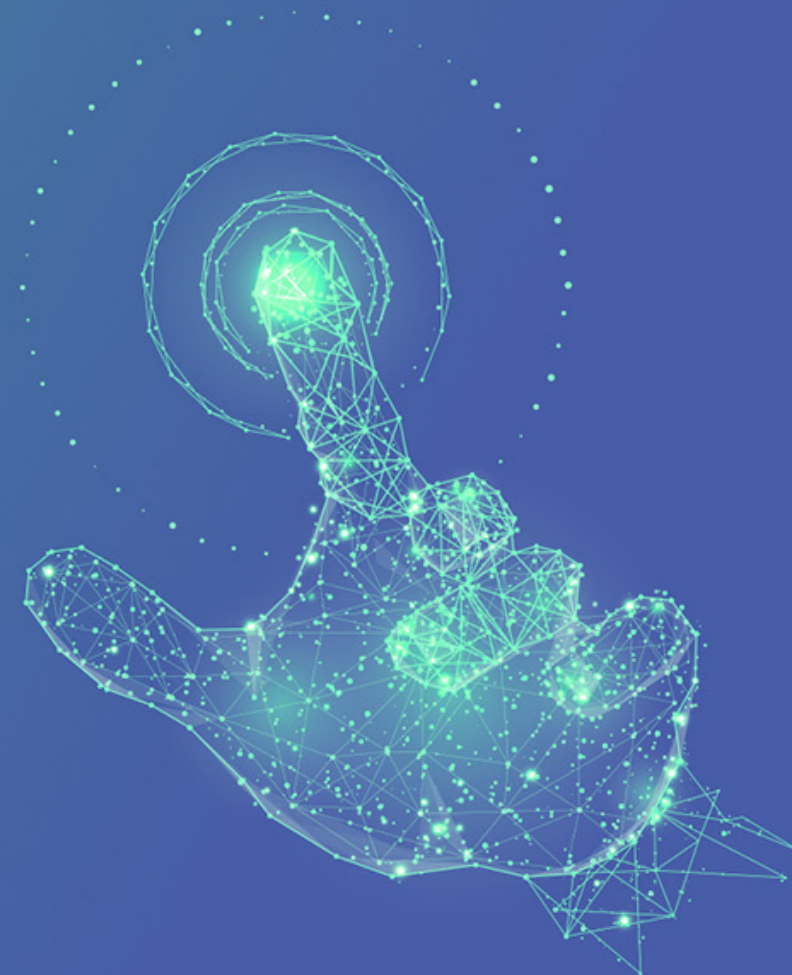
- 1) Establishing the current level of Multi-morbidity in the UK
- 2) Identify key Multimorbidity Clusters
- 3) Key trajectories of Multimorbidity Development

Scale:

c12 million



Better, Faster and More Efficient Clinical Trials



Better, Faster and More Efficient Clinical Trials

Enabling cost-effective clinical trials which exceed the standards required by international regulators and deliver results that improve patient care

Requires digital & regulatory innovation, pre-competitive knowledge sharing, and broad partnerships

20/21 Priorities

- **Development of NHS DigiTrial infrastructure** (with initial focus on Feasibility Service)
- **Building partnerships** with patient groups, medical charities, NIHR CRN, HRA, MHRA, government, tech & pharma
- **Securing further funding** to enable development of other aspects of NHS DigiTrial vision

QQR Deliverables

- **Development of new tools and methods** (e.g. for monitoring clinical trial quality; ascertaining clinical outcomes; and notifying trial participants of the results)
- **Demonstrator trials:** delivered in partnership with academic & commercial organizations; patient & researcher stories; metrics of researcher usage and patient access to trials
- **Changes to regulatory approaches:** new trials regulatory, ethics & information governance policies & practices

Addressing major challenges to quality & efficiency

Good design

Feasibility

- Assess number & location of potential participants
- Iteration of protocol & recruitment approach
- Focus efforts & resources

ORION-4

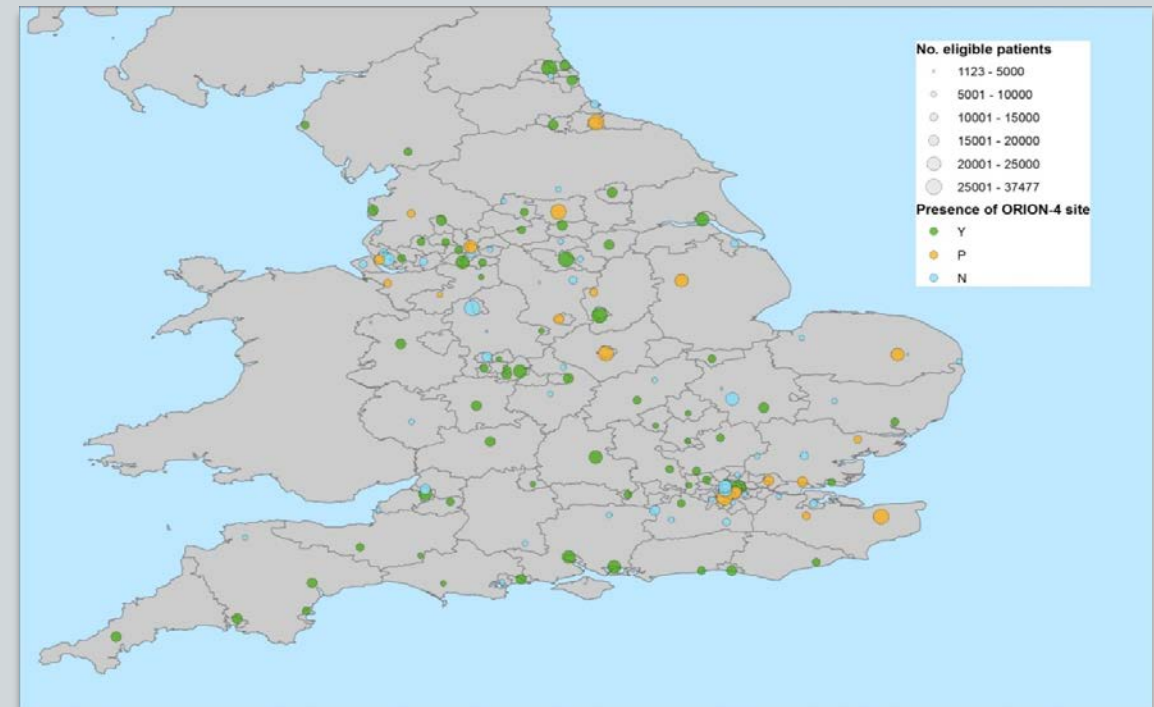
1.5 million people with cardiovascular disease at top 90 hospitals (over 10,000 per hospital) across whole of England

Recruitment ongoing
300,000 invitations

Up to 1000 phone calls/day

Typically >100 pts attending per hospital

(biggest sites district general hospitals along South Coast, Midlands, and North East)





News announcement

Wellcome to fund Health Data Research UK and The Alan Turing Institute PhD Programme

16 September 2019

Health Data Research UK, in partnership with The Alan Turing Institute, is to receive funding to support a PhD programme in Health Data Science funded by the Wellcome Trust.

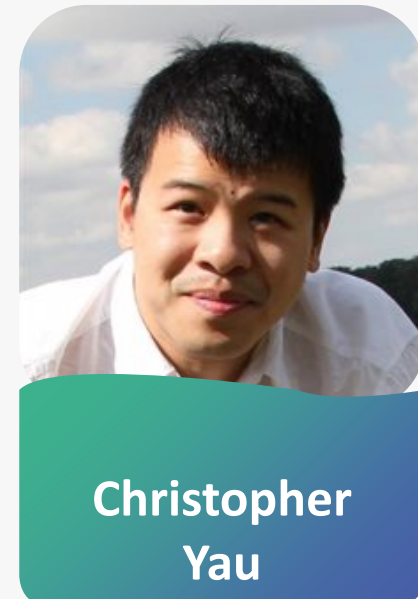
Researchers of the Future! Wellcome PhD Programme

HDRUK
Health Data Research UK

**The
Alan Turing
Institute**



Peter Diggle



**Christopher
Yau**

Strategic partnerships

British Heart Foundation

World-leading **BHF Data Science Hub** that capitalises on the UK's unique data assets to drive research and improve prevention, prediction and treatment of cardiovascular diseases and enhance delivery of care.

- Partnership approved - July 2018
- Submission to BHF April 2019
- Competitive call for Director
- November 2019 launch



Professor Sir Nilesh Samani



Professor Cathie Sudlow

This is just the start. We can achieve a lot more together....



**“Alone we can do so little;
Together we can do so much.”**
Helen Keller

Find out more

[About](#) [Science](#) [Talent & Training](#) [Digital Innovation Hubs](#) [News & Events](#) [Opinion](#)

What is health data science?



Health Data Research UK

We are uniting the UK's health data to make discoveries that improve people's lives. By making health data available to researchers and innovators we can better understand diseases and find ways to prevent, treat and cure them.

[More about us](#)

For further information:

 www.hdruk.ac.uk

 enquiries@hdruk.ac.uk

 [@HDR_UK](https://twitter.com/HDR_UK)