Stimulus Materials

Workshops for Stage B

Final

Stimulus 1a

We would like to create a resource to explain Secure Data Environments/ Trusted Research Environments to the public

This resource needs to cover these topics:

How they work

Who accesses the information in it

What the information is used for

Key relevant terms:
- Data access
- Data federation

What are the benefits

Security and privacy of information

Stimulus 1b

The resource is likely to have different layers of information

- A 1-line summary about Secure Data Environments/ Trusted Research Environments
- The main resource with further information
- A link to more detail for those who want it

Today, we need to think about..

- How we explain the topics on the previous page
- Which ‘layer’ of information each topic fits within
- What format this resource should be in

Stimulus 1c

What we learnt in the last workshops

1. Different people would like different amounts of information; particularly in relation to the security and privacy of information.
2. In this resource, it will be <u>important to communicate these points quickly</u>;
    - Whether data is identifiable
    - Who will access the data and how will they be approved
    - That the data is safely stored
    - What the purpose of accessing the data is – what are the benefits to patients/ members of the public
3. The terms ‘health data’, ‘data access’ and ‘approved users/ researchers’ were largely understood
4. Other terms required further explanation including: ‘de-identified’, ‘remote access’ and ‘analytical tools’

For the purposes of clarity, like last time we will refer to this method of data storage and access as Secure Data Environments or Trusted Research Environments. These are both different phrases used to describe the same thing.

Stimulus 2 – One line summary with 4 alternatives

What are Secure Data Environments/ Trusted Research Environments?

a. Secure data storage and access environments that provide remote access to de-identified health data for approved individuals.
b. Give approved users access to health data that is not identifiable for analysis, without them needing to download the data.
c. A place where approved users, including researchers, NHS providers and service planners, and charities, can access de-identified health data for use in research that can save and improve lives.
d. A controlled and safe way of accessing health data.

How do they work?

Principles that **all Secure Data Environments/ Trusted Research Environments follow:**

a. Data access – data is accessed, rather than copied or shared.
b. Data cannot be taken out, only analysis/results.
c. Enables different sources of data to be linked together.

How do they work?

a. The data remains in a secure location, and approved individuals access it remotely. Individual level data cannot be taken out; they can only export analysis results (such as tables and figures), and only after careful checks have been made.

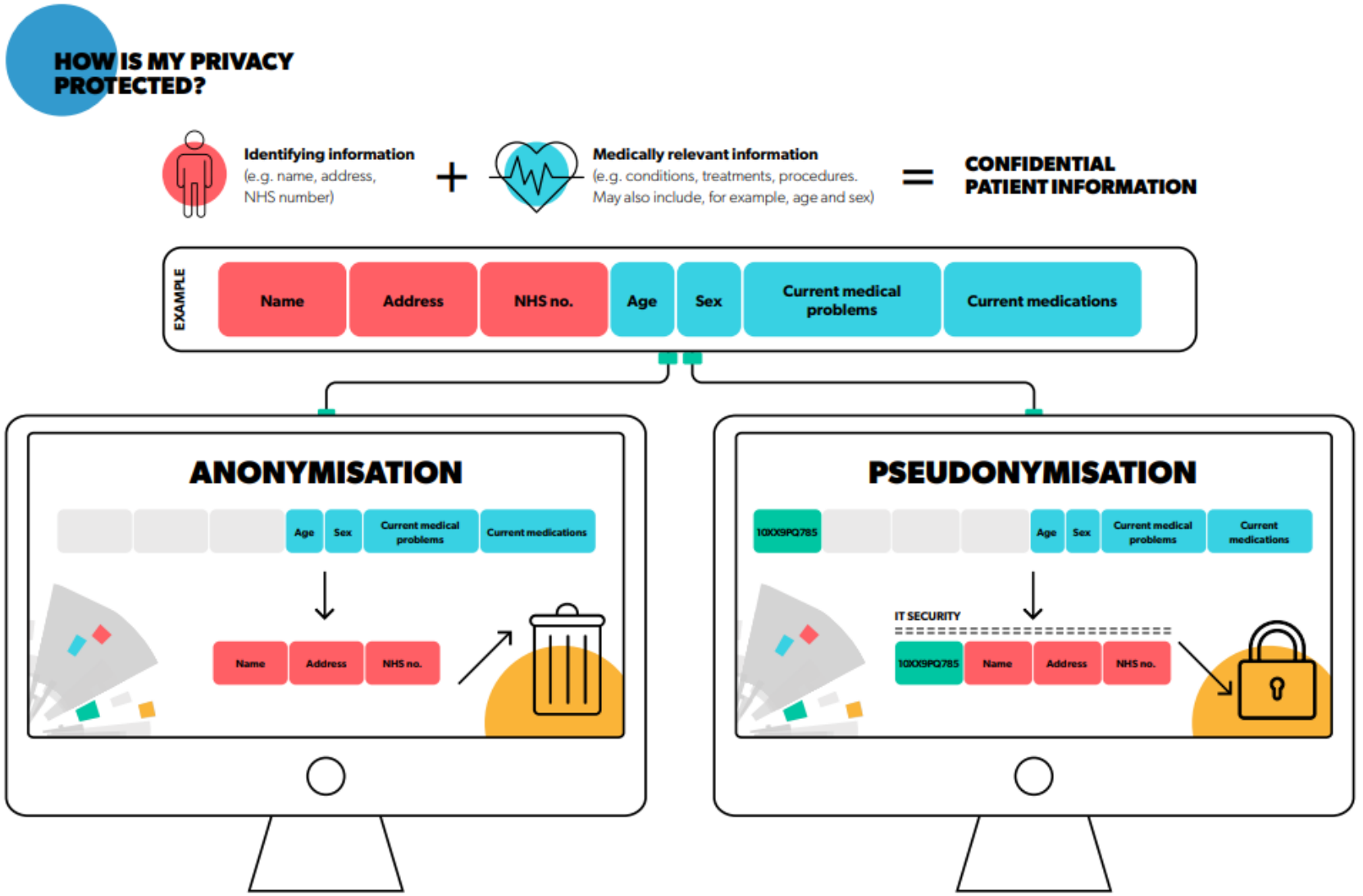
b. Data is housed in secure environments. Researchers then apply for remote access to this data. This means that instead of multiple copies of a data set being sent out to multiple researchers, researchers can access the data they need in a single secure location.

How do they work?

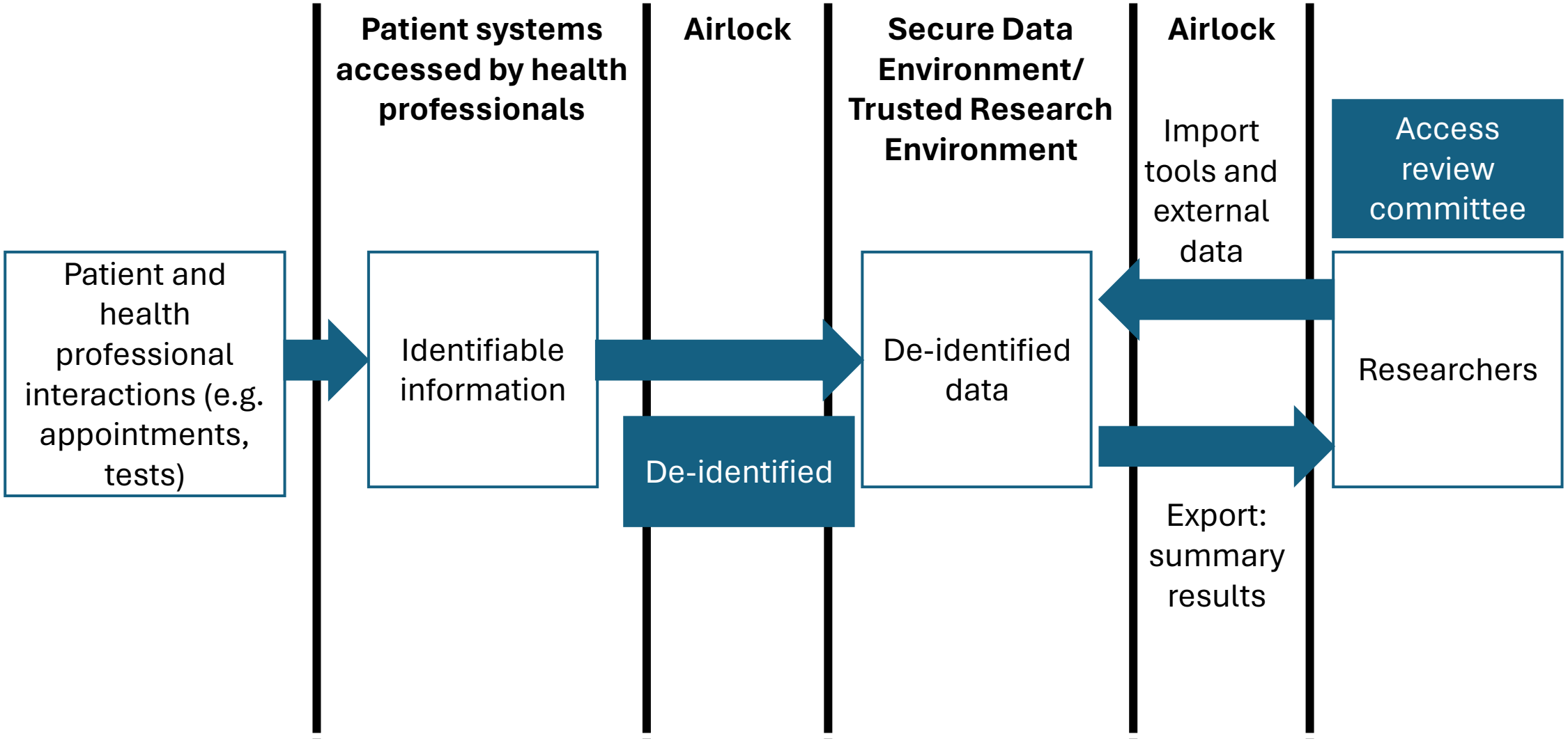
a. Some of them involve data storage and access in one place and others involve federation across platforms.
   - Examples that use federated data: NHS England Federated Data Platform, Open Safely and Public Health Scotland's National safe haven.
   - Examples that involve data storage: NHS England Secure Data Environment, HSC Northern Ireland's Honest Broker Service TRE, Public Health Scotland's National Safe Haven.

b. Some use de-identified data, others use identifiable data. In both cases robust processes are in place to maintain patient confidentiality.
   - For example, the NHS England Federated data platform uses identifiable information (for the use of clinicians in direct care of patients)
   - Examples that use de-identified data are:
     - NHS England Secure Data Environment
     - Public Health Scotland's National Safe Haven
     - HSC Northern Ireland's Honest Broker Service TRE
     - SAIL databank (Wales)
     - Discover-NOW – a London Secure Data Environment

Stimulus 3d – De-identification (if needed)



Diagram


Patient systems accessed by health professionals
Airlock
Secure Data Environment/ Trusted Research Environment
Airlock
Patient and health professional interactions (e.g. appointments, tests)
Identifiable information
De-identified
De-identified data
Import tools and external data
Access review committee
Researchers
Export: summary results

How do they work – description of data federation

a. **Data federation** is not a data collection; it is software that helps to connect disparate sets of data and allow them to be used more effectively, e.g. for care.

b. **Data federation** is software that sits across organisations like NHS trusts and Integrated Care Systems allowing them to connect data they already hold in a secure and safe environment.*

c. **Data federation** involves the creation of a virtual database that maps many different sources and makes them accessible through a single interface. Unlike other technologies, data federation leaves all data at the source.**

Who accesses <u>de-identified</u> information

a. List of who may access it
    a. ‘Approved individuals’
    b. ‘Approved researchers’
    c. Researchers within and outside of the NHS.
    d. NHS providers and commissioners, university researchers, charities, companies such as private healthcare providers or pharmaceutical companies.*

How they get approval

Process they must go through to get approved

a. Approved by an access review committee.

b. To be considered Safe people, researchers have to: demonstrate that they have the technical skills to use the data, either through academic qualifications or practical research experience; complete our training course and pass the assessment at the end; agree to their details being published on our website; and sign an agreement promising to protect the confidentiality of your data at all times. *

c. For access to data to be granted, the researchers need to demonstrate that their proposal is an appropriate and ethical use of the data, that it will deliver clear public benefits and that they will publish their results to enable use, scrutiny and further research.*

d. When patient data is used for purposes beyond providing individual care, there are safeguards to ensure it is protected and used appropriately. Organisations follow set processes and criteria for the decisions they make about how patient data can be used, some of which are set out in law. There are also groups and committees who are often independent from these organisations involved in the process. It’s their job to give advice to these organisations on what would or would not be lawful and appropriate uses of patient data**

Benefits of Secure Data Environments/ Trusted Research Environments

a. Secure Data Environments/ Trusted Research Environments help make research efficient, collaborative and cost effective, providing rich data that enables deep insights which will go on to improve healthcare and save lives.

b. People can be confident that their personal health data is or will be accessed securely and their privacy is always protected.

c. Access to data can be recorded to track who has had access to what data.

d. Specific examples of outcome benefits:

- NHS England Secure Data environment: provides academics with access to cardiovascular and cancer data for Covid-19 research.
- The 1000 genomes project: analyse genomic and long-term clinical data (from health records) to gain insight into the nature of genetic changes that drive cancer evolution. (An opt-in study)
- SAIL databank in Wales: Linking police and healthcare data has revealed that highly vulnerable individuals are detectable before the involvement of the police. Sharing data can allow for efficient, targeted allocation of resources – both for police by preventing future callouts and for healthcare by preventing visits from injury, reducing the strain on the NHS.
- Office of National Statistics: research explored the representation and career progression of people from different ethnic backgrounds within the teaching profession in England.

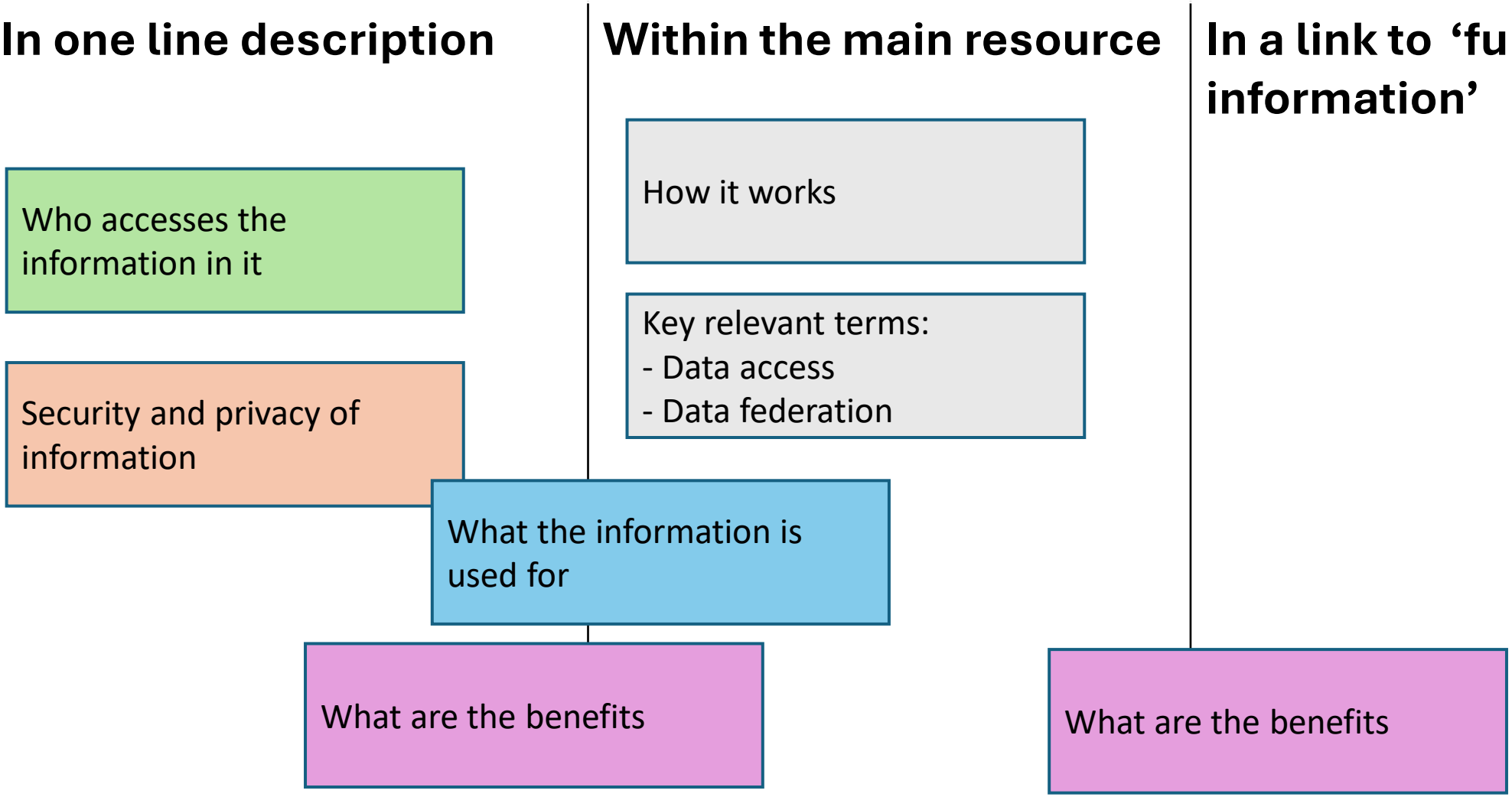
Uses – alternative levels of detail (using NHSE as e.g.)

a. Used for projects in the public interest.
b. All uses of the data must be for the public good.
c. For projects that will deliver clear public benefits.
d. To save and improve lives.
e. Providing academics with access to cardiovascular and cancer data for Covid-19 research.
f. For secondary uses such as research and analysis.
g. To identify better ways to predict and diagnose illness, to develop new treatments and monitor the safety of existing treatments, for planning services and for addressing health inequalities.

Security messages

a. Safe people: only approved researchers are granted access to the data.
b. Safe projects: data is only made accessible for projects in the public interest.
c. Safe settings: Data is accessed by researchers in a secure room, or via an approved, remote connection to one.
d. Safe data: Data is de-identified and only data that is really needed for the project is made accessible.
e. Safe output: All research findings are checked by staff with any potentially identifying information removed.

Stimulus 8 – prioritising the information



Part 2 – what format?

Stimulus 9

The main resource could be produced in a variety of formats

- An animated video.
- An infographic – visual 1-page sheet of information.
- A web-page.

Stimulus 10a

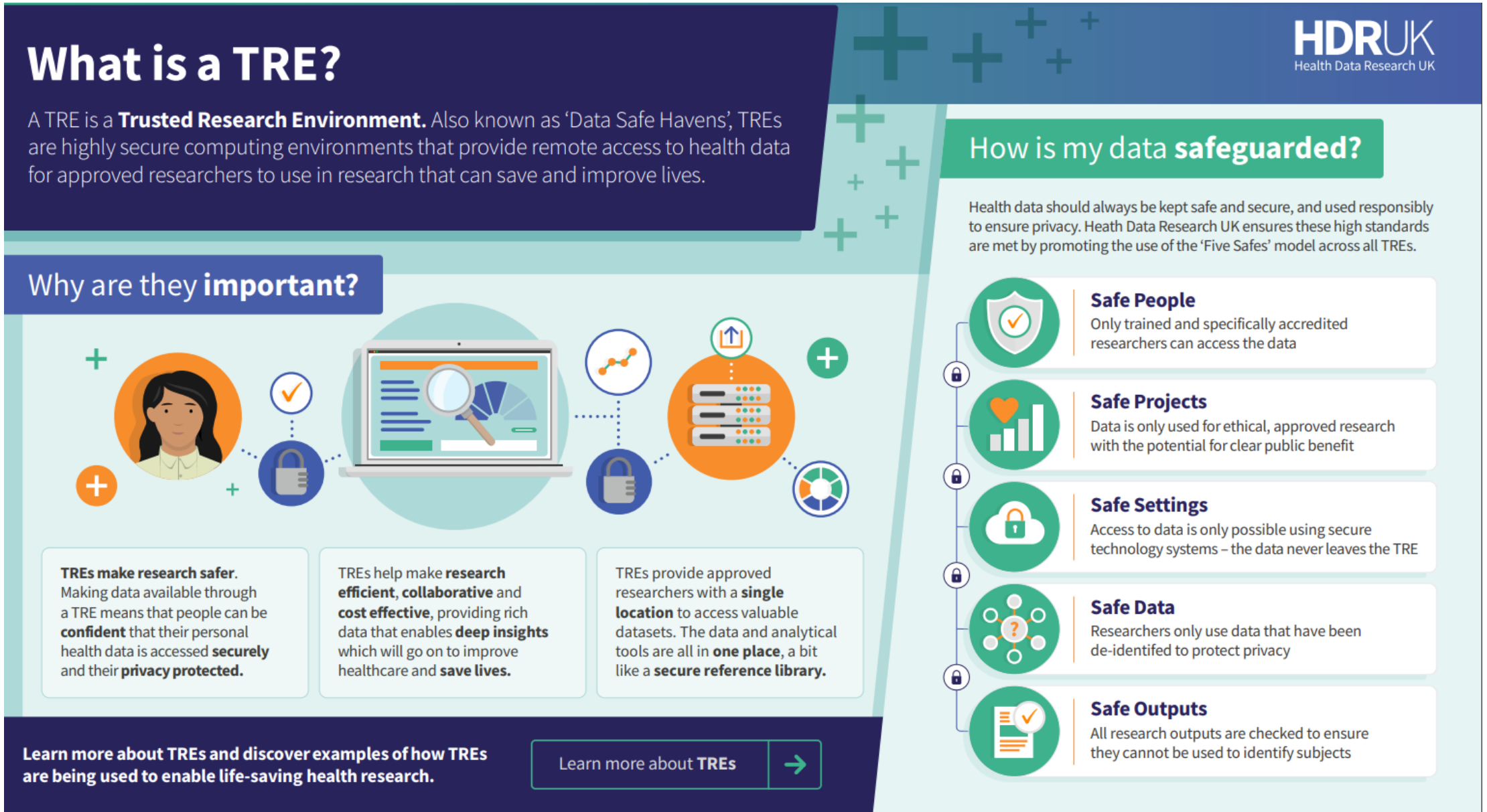
Example animations

Introducing patient data | Understanding patient data

A health data journey (youtube.com)

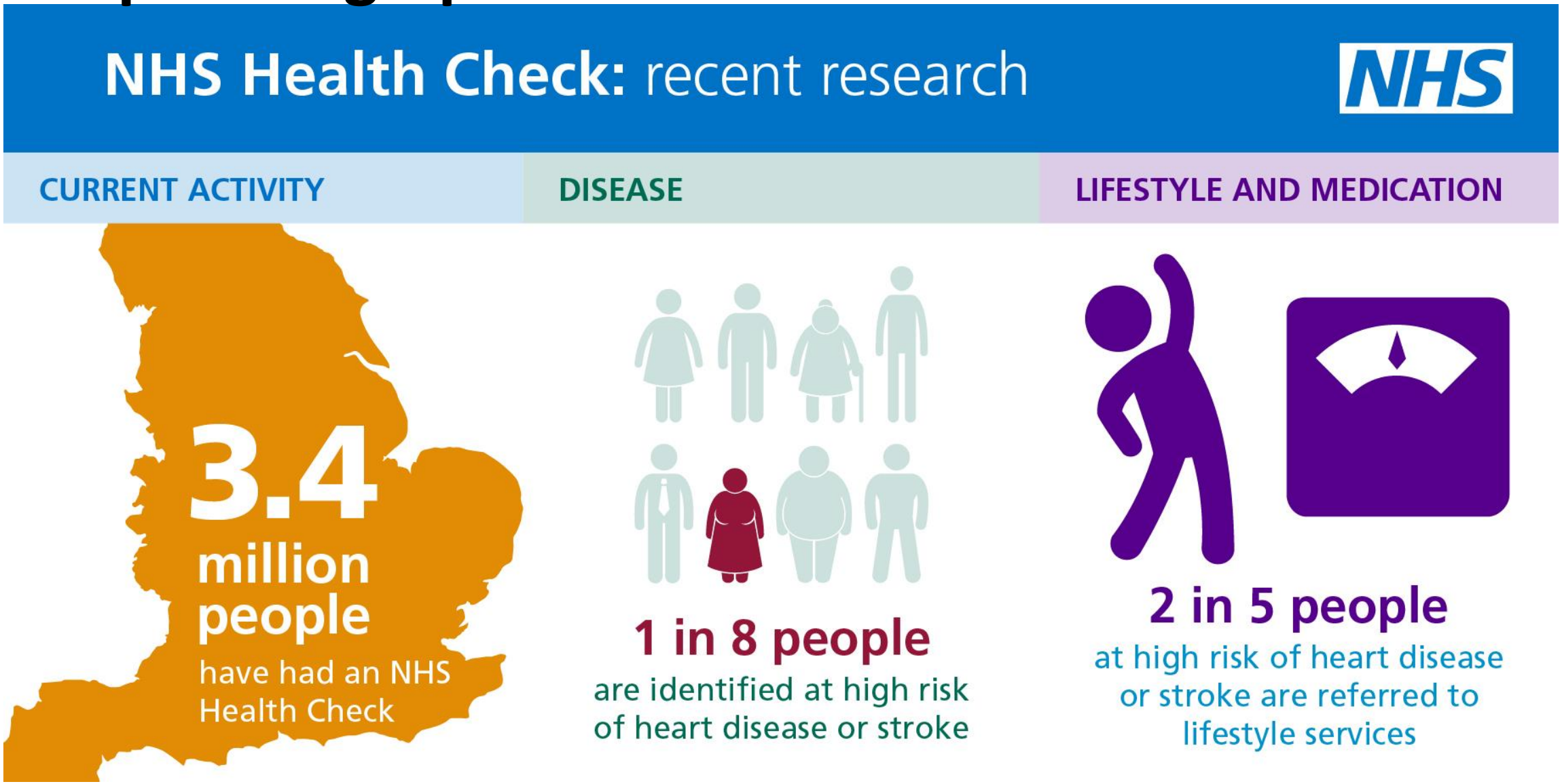
Stimulus 10b

Example infographic



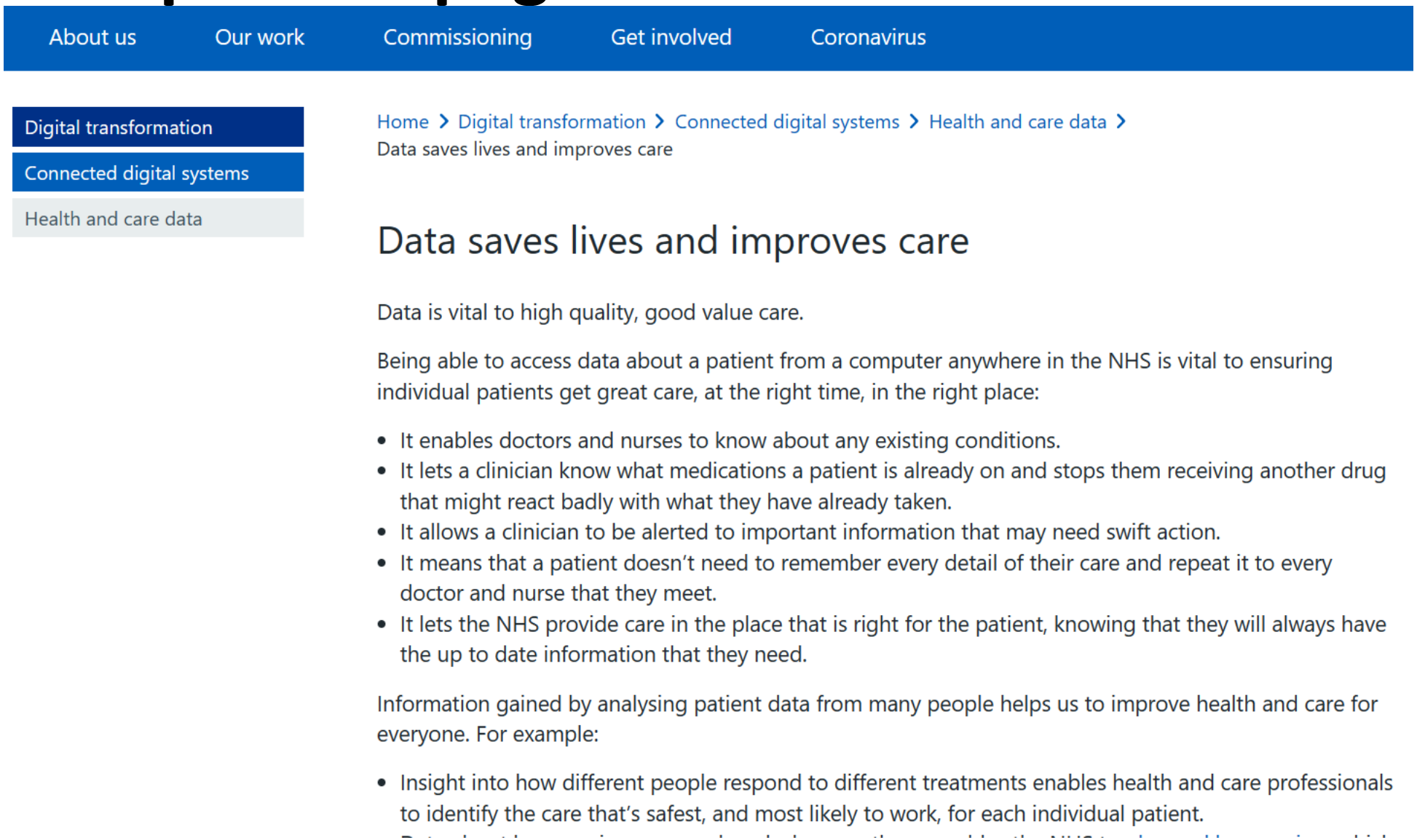
Stimulus 10b

Example infographic



Stimulus 10c

Example web page



Part 3 – co-create

Stimulus 11

Write the script/design the 1 pager