

# Question driven harmonisation of data The variable cascade in practice

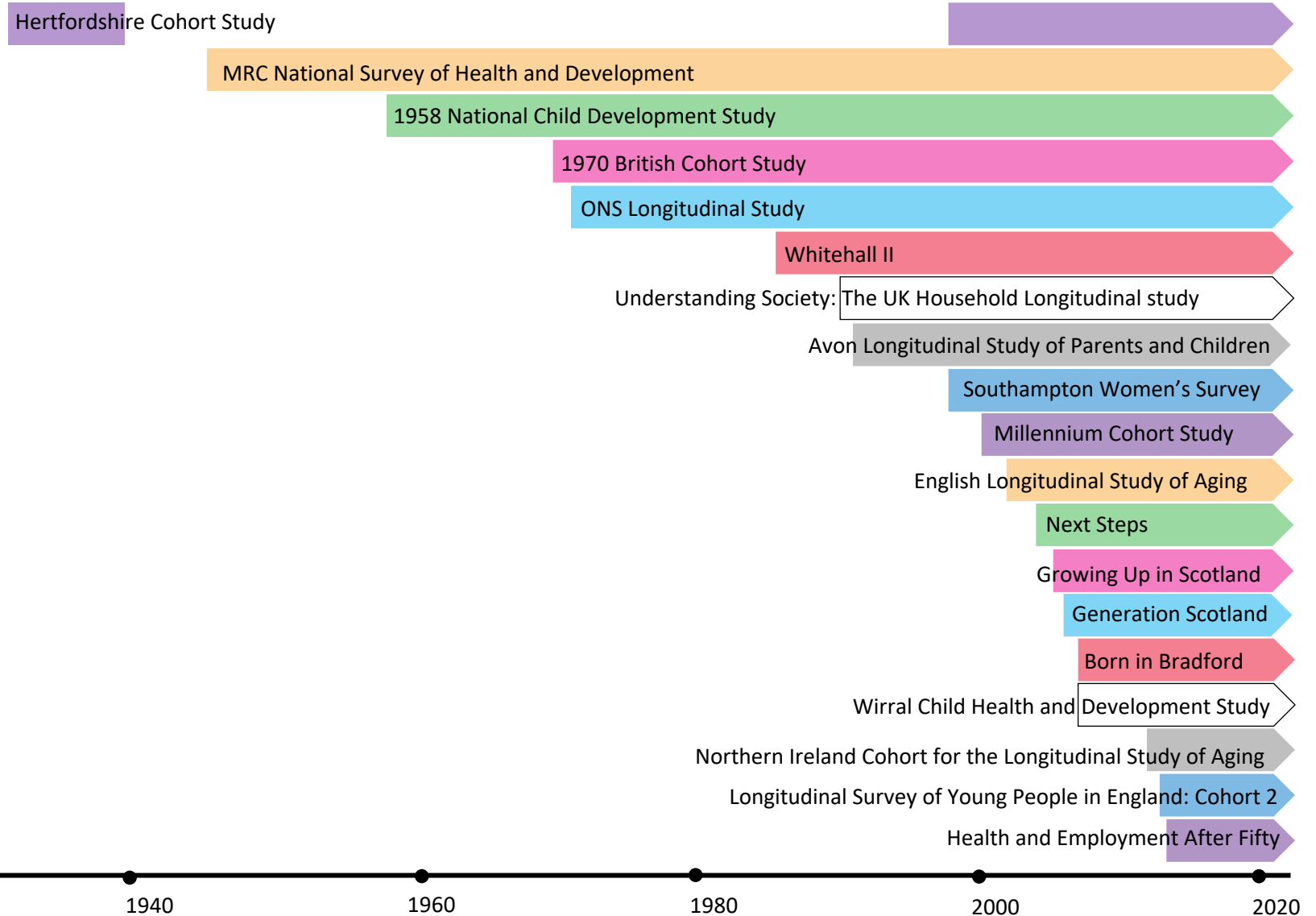
Jon Johnson & Bill Norley

CLOSER, UCL Social Research Institute

December 2020


License: [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)

# Partners




The picture can't be displayed.


The picture can't be displayed.



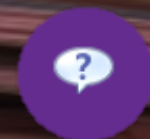
10  
Studies



387  
Datasets



313  
Questionnaires



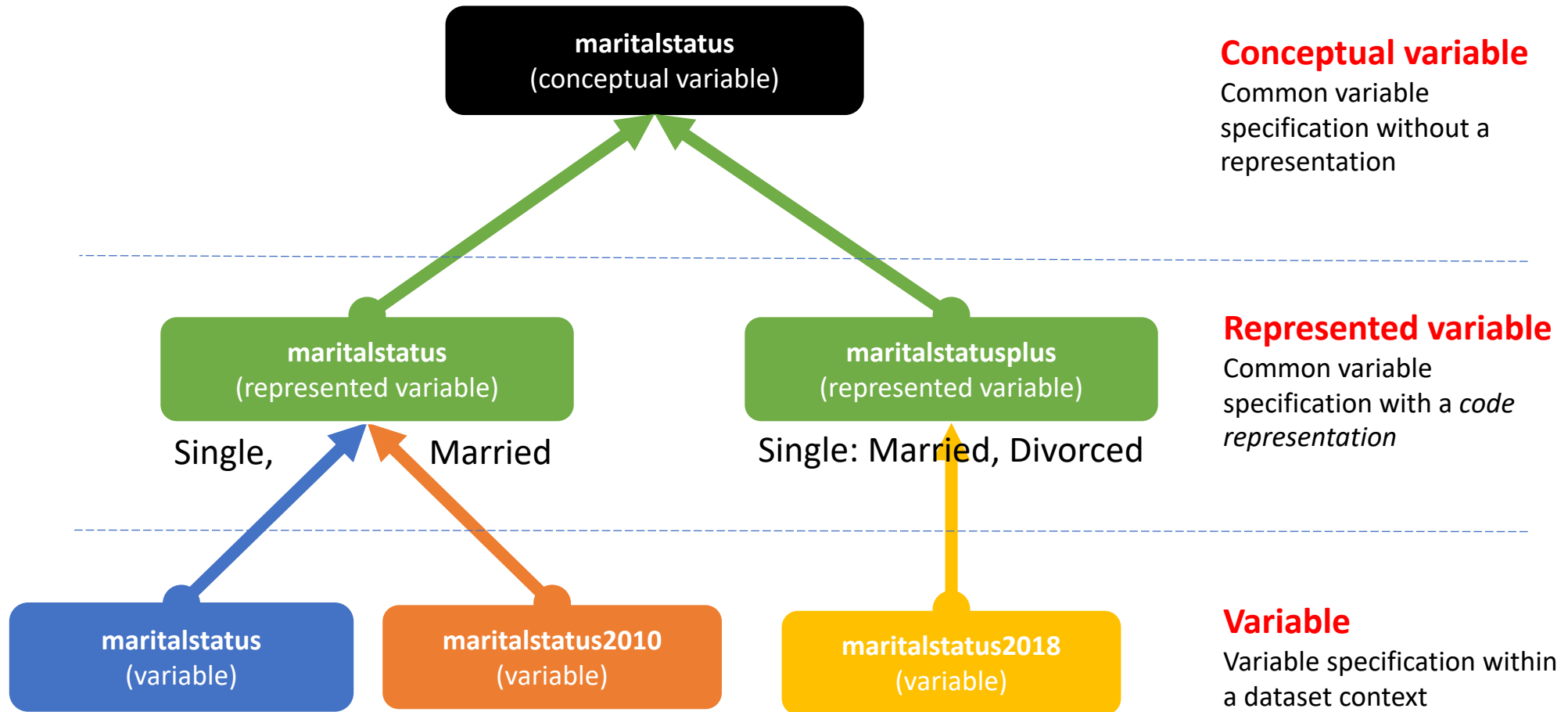
43,490  
Questions



144,080  
Variables

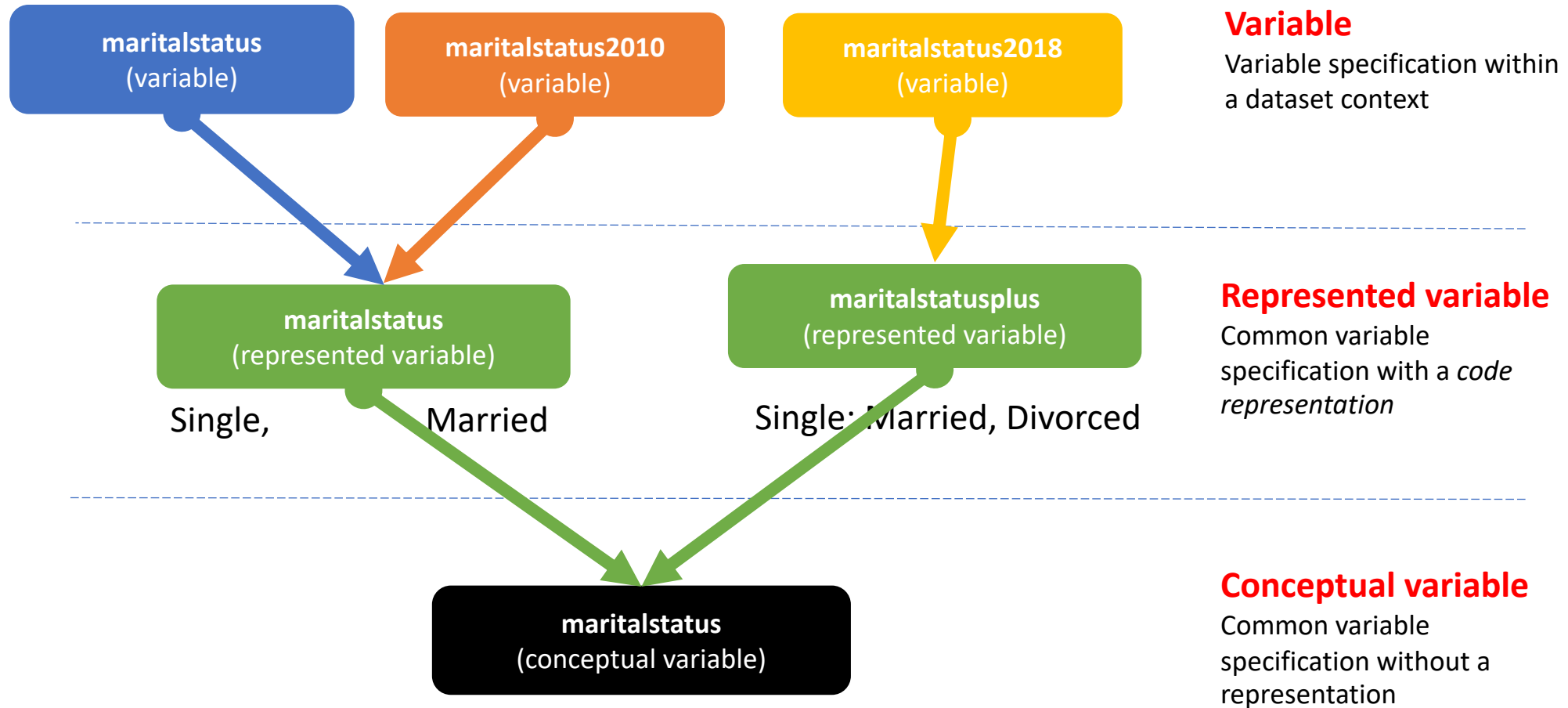


# Documenting comparabilities among variables



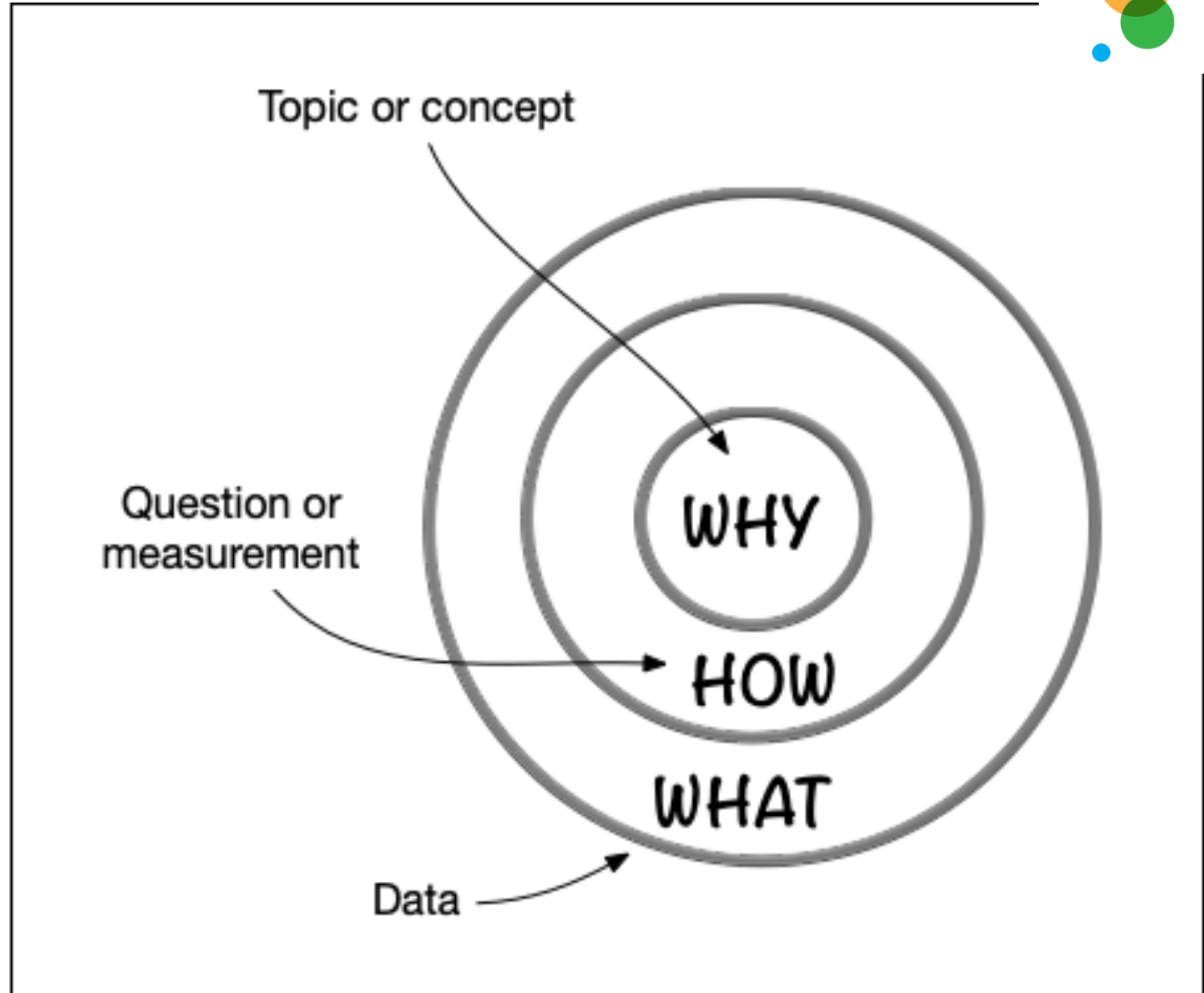


# Organising 144,000 variables



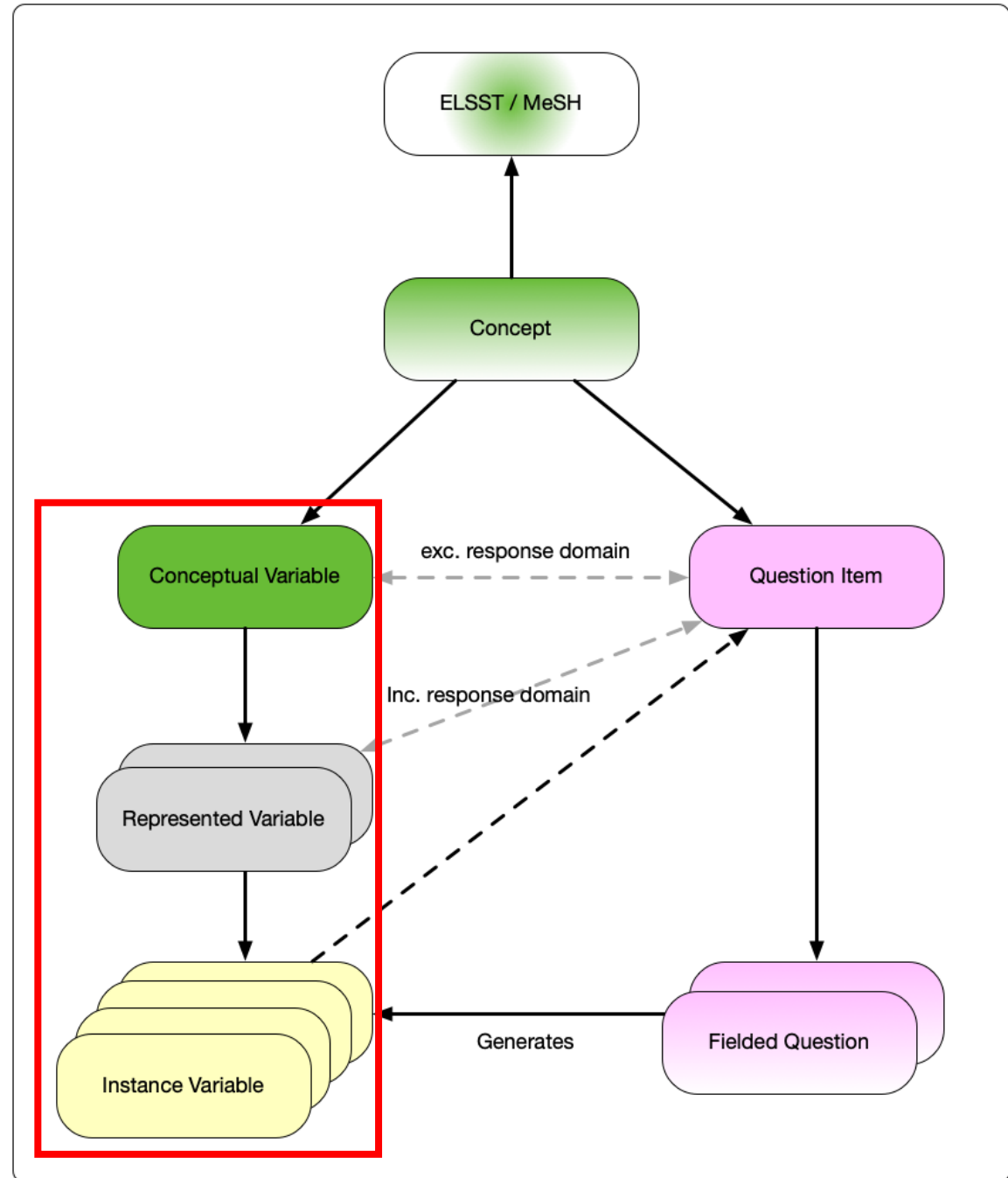


Questions or measurements capture the concept and result in data



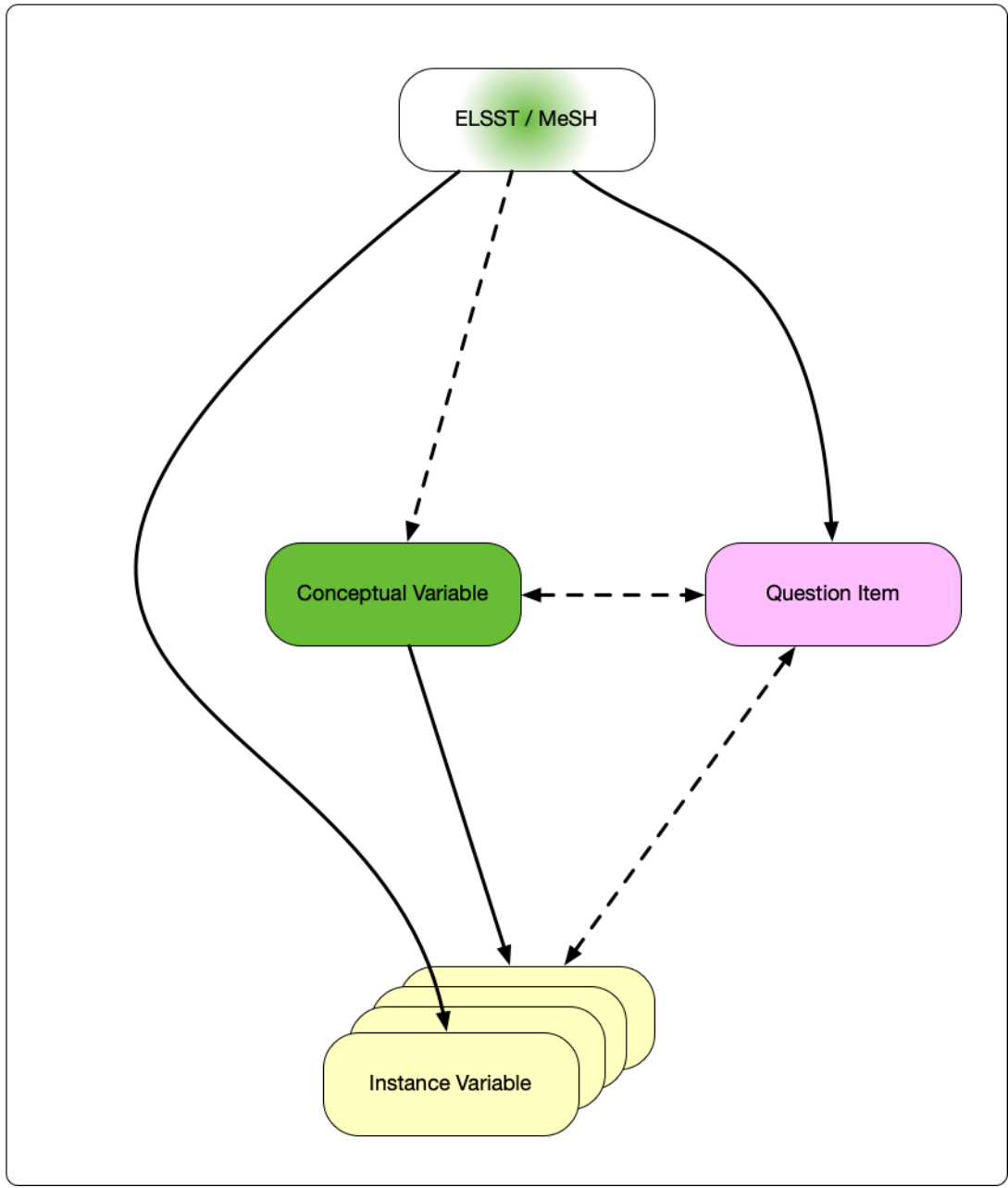


Variable  
Cascade  
does not sit  
in a vacuum





# Conceptual Variables driven by questions





We are not  
creating  
represented  
variables  
(yet)

The aim of CLOSER Discovery is to provide information to allow researchers to make their own decisions e.g. for harmonisation

We are not the studies who may (or may not) have views about what they would recommend

It might be that some users would find represented variables useful, for some it might not be the case.

Its also a lot of work

# Managing Conceptual Variables

## DDI-Lifecycle Colectica Repository

- High level topics e.g. ELSST or MeSH
- Question information (question name, question text)
- Variable information (variable name and label)

## Colectica SDK

# Expected Output

```
[{"text":"COVID-19 Wave 1","id":"uk.iser 3f603afc-c53a-4609-9b94-ec238b2b8e8a","parent":"uk.iser"}, {"text":"COVID-19 Wave 2","id":"uk.iser 9b054291-6aa6-4cd4-b4ad-1863b8a1aa0f","parent":"uk.iser"}, {"text":"COVID-19 Wave 3","id":"uk.iser 3b0f3340-5ce9-4f48-b56f-4f5fb9556c3e","parent":"uk.iser"}]
```

- 1970 British Cohort Study
- Avon Longitudinal Study of Parents and Children
- Hertfordshire Cohort Study
- Millennium Cohort Study
- National Child Development Study
- National Survey of Health and Development
- Next Steps
- Southampton Women's Survey
- Understanding Society
- Wirral Child Health and Development Study

## Index - AllItems = 1648

There are a number of methods for matching equivalences in Studies. The methods below will only provide a match where there is 100% correlation.

Match Text  
 Match All Words  
 Match Name  Match Text

Please select the equivalences to be analysed and then sweep(s) to be used with these equivalences

Select a file of Equivalences to Upload ...

No file selected.

or Add word(s) to search. You can add or delete word(s) to search by using the appropriate buttons.

currently living partner

childcare home schooling

experiencing symptoms coronavirus

currently pregnant

vouchers meals schools

work at home

unhappy depressed

# Expected Output

## Identifying Equivalent Question - AllItems = 1648

```
[[{"text": "COVID-19 Wave 1", "id": "uk.iser 3f603afc-c53a-4609-9b94-ec238b2b8e8a", "parent": "uk.iser"}, {"text": "COVID-19 Wave 2", "id": "uk.iser 9b054291-6aa6-4cd4-b4ad-1863b8a1aa0f", "parent": "uk.iser"}, {"text": "COVID-19 Wave 3", "id": "uk.iser 3b0f3340-5ce9-4f48-b56f-4f5fb9556c3e", "parent": "uk.iser"}]]
```

| ID | Search Term              | Question Name   | Question Text  | Variable Name     | Variable Text                                | Agency  | Study           |
|----|--------------------------|---|--|-------------------|--|---------|-----------------|
| 1  | currently living partner | <input checked="" type="checkbox"/> qi_couple         | Are you currently living with a partner?   | ca_couple         | Living with a partner                        | uk.iser | COVID-19 Wave 1 |
|    |                          | <input checked="" type="checkbox"/> qi_couplewsh      | Are you currently living with a partner?   | cb_couplewsh      | Living with a partner                        | uk.iser | COVID-19 Wave 2 |
|    |                          | <input checked="" type="checkbox"/> qi_couplewsh      | Are you currently living with a partner?   | cc_couplewsh      | Living with a partner                        | uk.iser | COVID-19 Wave 3 |
| 2  | childcare home schooling | <input checked="" type="checkbox"/> qi_timechcare     | About how many hours did you spend on childcare or home schooling last week? Hours per week  | ca_timechcare     | Time spent on childcare or home schooling    | uk.iser | COVID-19 Wave 1 |
|    |                          | <input checked="" type="checkbox"/> qi_timechcare     | About how many hours did you spend on childcare or home schooling last week? ... Hours per week  | cb_timechcare     | Time spent on childcare or home schooling    | uk.iser | COVID-19 Wave 2 |
|    |                          | <input checked="" type="checkbox"/> qi_hrchs          | Have you reduced the hours you spend in your paid job because of the time you spend doing childcare or home schooling?   | cb_hrchs          | Employment hours change due to homeschooling | uk.iser | COVID-19 Wave 2 |
|    |                          | <input checked="" type="checkbox"/> qi_workchs        | Have you adapted the work schedule of your paid job because of the time you spend doing childcare or home schooling?   | cb_workchs        | Work pattern change due to homeschooling     | uk.iser | COVID-19 Wave 2 |
|    |                          | <input checked="" type="checkbox"/> qi_workchs        | Have you adapted the work schedule of your paid job because of the time you spend doing childcare or home schooling?   | cc_workchs        | Work pattern change due to homeschooling     | uk.iser | COVID-19 Wave 3 |
|    |                          | <input checked="" type="checkbox"/> qi_timechcare     | About how many hours did you spend on childcare or home schooling last week? ... Hours per week  | cc_timechcare     | Time spent on childcare or home schooling    | uk.iser | COVID-19 Wave 3 |
|    |                          | <input checked="" type="checkbox"/> qi_hrchs          | Have you reduced the hours you spend in your paid job because of the time you spend doing childcare or home schooling?   | cc_hrchs          | Employment hours change due to homeschooling | uk.iser | COVID-19 Wave 3 |
| 3  | currently pregnant       | <input checked="" type="checkbox"/> qi_preg           | Are you currently pregnant?  | ca_pregnow        | Whether pregnant                             | uk.iser | COVID-19 Wave 1 |
|    |                          | <input checked="" type="checkbox"/> qi_pregnow        | Are you currently pregnant?  | cb_pregnow        | Whether pregnant                             | uk.iser | COVID-19 Wave 2 |
|    |                          | <input checked="" type="checkbox"/> qi_pregnow        | Are you currently pregnant?  | cc_pregnow        | Whether pregnant                             | uk.iser | COVID-19 Wave 3 |
| 4  | work at home             | <input checked="" type="checkbox"/> qi_hhearn_period  | Thinking about everyone living with you at the moment, what is the total take-home pay/earnings of your household now? Please only include earnings from paid work or self-employment. If you are not sure, please ca_hhearn_period tell us an approximate amount. Per   |                   | Current household earnings period            | uk.iser | COVID-19 Wave 1 |
|    |                          | <input checked="" type="checkbox"/> qi_hhearn_amount  | Thinking about everyone living with you at the moment, what is the total take-home pay/earnings of your household now? Please only include earnings from paid work or self-employment. If you are not sure, please ca_hhearn_amount tell us an approximate amount. ... Pounds  |                   | Current household earnings amount            | uk.iser | COVID-19 Wave 1 |
|    |                          | <input checked="" type="checkbox"/> qi_blhearn_period | Thinking about everyone who was living with you in January/February 2020, what was the usual total take-home pay/earnings of your household in January/February 2020? Please only include earnings from paid work or self-employment, after tax, National Insurance and pension contributions have been deducted. If you are not sure, please tell us an approximate amount. Per | ca_blhearn_period | Baseline: household earnings period          | uk.iser | COVID-19 Wave 1 |



# Expected Output

## Identifying Equivalent - AllItems = 1648

```
[{"text": "COVID-19 Wave 1", "id": "uk.iser.3f603afc-c53a-4609-9b94-ec238b2b8e8a", "parent": "uk.iser"}, {"text": "COVID-19 Wave 2", "id": "uk.iser.9b054291-6aa6-4cd4-b4ad-1863b8a1aa0f", "parent": "uk.iser"}, {"text": "COVID-19 Wave 3", "id": "uk.iser.3b0f3340-5ce9-4f48-b56f-4f5fb9556c3e", "parent": "uk.iser"}]
```

| ID | Search Term              | Name             | Text                                      | Variable Name      | Variable Text                                | Variable Item  | Question Name      | Question Text  |
|----|--------------------------|------------------|---|--------------------|--|--|--------------------|--|
| 1  | currently living partner | ca_couple        | Living with a partner                     | ca_couple          | Living with a partner                        | uk.iser.ukhls:7113e2c5-64b6-4f36-bf49-2f5f88f27359:3 | qi_couple          | Are you currently living   |
|    |                          |                  |   | cb_couplewsh       | Living with a partner                        | uk.iser.ukhls:12d52199-3699-4286-ad34-3ba175665c73:3 | qi_couplewsh       | Are you currently living   |
|    |                          |                  |   | cc_couplewsh       | Living with a partner                        | uk.iser.ukhls:cdc9ea7c-d0ea-418f-879f-c217508f906d:3 | qi_couplewsh       | Are you currently living   |
| 2  | childcare home schooling | ca_timechcare    | Time spent on childcare or home schooling | ca_timechcare      | Time spent on childcare or home schooling    | uk.iser.ukhls:6e684650-c3c2-4735-a99f-71e20d3ed3ef:3 | qi_timechcare      | About how many hours   |
|    |                          |                  |   | cb_timechcare      | Time spent on childcare or home schooling    | uk.iser.ukhls:c5eb7742-c311-4bd1-8dd2-df058273e823:3 | qi_timechcare      | About how many hours   |
|    |                          |                  |   | cb_hrchs           | Employment hours change due to homeschooling | uk.iser.ukhls:a16d4034-a103-4987-9bd5-8e5617134eac:3 | qi_hrchs           | Have you reduced the hours of schooling?   |
|    |                          |                  |   | cb_worksch         | Work pattern change due to homeschooling     | uk.iser.ukhls:16e88ee2-d564-4cb3-9e7b-74a8e44577be:3 | qi_worksch         | Have you adapted the way you are working at home?  |
|    |                          |                  |   | cc_worksch         | Work pattern change due to homeschooling     | uk.iser.ukhls:6059e6b6-1bd3-41f6-b70f-7c44ad8d461b:3 | qi_worksch         | Have you adapted the way you are working at home?  |
|    |                          |                  |   | cc_timechcare      | Time spent on childcare or home schooling    | uk.iser.ukhls:1caa1281-fadc-40c3-a62c-46fc0638a441:3 | qi_timechcare      | About how many hours   |
|    |                          |                  |   | cc_hrchs           | Employment hours change due to homeschooling | uk.iser.ukhls:64cec6cb-1aff-4629-8b60-e0bd0b75a482:3 | qi_hrchs           | Have you reduced the hours of schooling?   |
| 3  | currently pregnant       | ca_pregnow       | Whether pregnant                          | ca_pregnow         | Whether pregnant                             | uk.iser.ukhls:5e311454-ae63-47e6-b0e6-a6410689bb4f:3 | qi_preg            | Are you currently pregnant?  |
|    |                          |                  |   | cb_pregnow         | Whether pregnant                             | uk.iser.ukhls:8b191d72-76f0-412b-b4f6-aed44c7d578c:3 | qi_pregnow         | Are you currently pregnant?  |
|    |                          |                  |   | cc_pregnow         | Whether pregnant                             | uk.iser.ukhls:6a498ba2-6bec-43e4-8a52-4b0b489a09e3:3 | qi_pregnow         | Are you currently pregnant?  |
| 4  | work at home             | ca_hhearn_period | Current household earnings period         | ca_hhearn_period   | Current household earnings period            | uk.iser.ukhls:78206333-a4f4-4cfa-a32d-34bfd1d3f326:3 | qi_hhearn_period   | Thinking about everyone in your household now? Please use an approximate amount.                           |
|    |                          |                  |   | ca_hhearn_amount   | Current household earnings amount            | uk.iser.ukhls:cb52b378-0625-4be1-84c9-00bb7d04df2e:2 | qi_hhearn_amount   | Thinking about everyone in your household now? Please use an approximate amount.                           |
|    |                          |                  |   | ca_blhhearn_period | Baseline: household earnings period          | uk.iser.ukhls:4cf6b8b9-38b3-4600-bc37-f2e53e261b13:3 | qi_blhhearn_period | Thinking about everyone's pay/earnings of your household, after tax, please tell us an approximate amount. |
|    |                          |                  |   | ca_wah             | Working at home                              | uk.iser.ukhls:cfe01df0-2f77-4f26-8ad9-ae0cc0e745ea:3 | qi_wah             | During the last four weeks, please tell us an approximate amount.  |
|    |                          |                  |   | ca_hstime          | Time taken - child care                      | uk.iser.ukhls:aa31f01a-0764-4cca-8629-08d1493a40bd:3 | oi_hstime          | Thinking about the situation...  |



- Demographics
- Housing and local environment
- Physical health
- Mental health and mental processes
- Health care
- Health behaviour
- Family and social networks
- Education
- Employment and income
- Expectations, attitudes and beliefs
- Child development
- Life events
- Omics
- Pregnancy
- Administration
- COVID-19
  - Demographics (CV19)
  - Housing and local environment (CV19)
  - Physical health (CV19)
  - Mental health and mental processes (CV19)
  - Health care (CV19)
  - Health behaviour (CV19)
  - Family and social networks (CV19)
  - Education (CV19)
  - Employment and income (CV19)
  - Expectation, attitudes and beliefs (CV19)
  - Child development (CV19)

|   | COVID-19 Wave 1 | COVID-19 Wave 2  | COVID-19 Wave 3  |
|---|-----------------|------------------|------------------|
|   | ca_indresp_w    | cb_indresp_w     | cc_indresp_w     |
| <b>+</b> New health conditions diagnosed - An emotional, nervous or psychiatric problem |                 | cb_hcondnew_cv22 | cc_hcondnew_cv22 |
| <b>+</b> New health conditions diagnosed - Angina                                       |                 | cb_hcondnew_cv5  | cc_hcondnew_cv5  |
| <b>+</b> New health conditions diagnosed - Any kind of liver condition                  |                 | cb_hcondnew_cv12 | cc_hcondnew_cv12 |
| <b>+</b> New health conditions diagnosed - Are very overweight                          |                 | cb_hcondnew_cv27 | cc_hcondnew_cv27 |
| <b>+</b> New health conditions diagnosed - Arthritis                                    |                 | cb_hcondnew_cv2  | cc_hcondnew_cv2  |
| <b>+</b> New health conditions diagnosed - Asthma                                       |                 | cb_hcondnew_cv1  | cc_hcondnew_cv1  |
| <b>+</b> New health conditions diagnosed - Cancer or                                    |                 | cb_hcondnew_cv13 | cc_hcondnew_cv13 |

- Demographics
- Housing and local environment
- Physical health
- Mental health and mental processes
- Health care
- Health behaviour
- Family and social networks
- Education
- Employment and income
- Expectations, attitudes and beliefs
- Child development
- Life events
- Omics
- Pregnancy
- Administration
- COVID-19
  - Demographics (CV19)
  - Housing and local environment (CV19)
  - Physical health (CV19)
  - Mental health and mental processes (CV19)
  - Health care (CV19)
  - Health behaviour (CV19)
  - Family and social networks (CV19)
  - Education (CV19)
  - Employment and income (CV19)
  - Expectation, attitudes and beliefs (CV19)
  - Child development (CV19)

|   |  |                    |
|---|--|--------------------|
| <p><b>+</b> Participant had loss of sense of smell or taste</p> | <p>covid1_2086<br/>covid1_2084<br/>covid1_2082<br/>covid1_2087<br/>covid1_2080<br/>covid1_2083<br/>covid1_2088<br/>covid1_2081<br/>covid1_2085</p> | <p>covid2_1008</p> |
| <p><b>+</b> Participant had nausea and/or vomiting</p>          | <p>covid1_2016<br/>covid1_2018<br/>covid1_2015<br/>covid1_2017<br/>covid1_2010<br/>covid1_2011<br/>covid1_2013<br/>covid1_2014<br/>covid1_2012</p> | <p>covid2_1001</p> |
| <p><b>+</b> Participant had numbness or tingling</p>            | <p>covid1_2216<br/>covid1_2211<br/>covid1_2214<br/>covid1_2212<br/>covid1_2213<br/>covid1_2218<br/>covid1_2210</p>                                 | <p>covid2_1021</p> |



# Summary



Matching criteria based on the actual question asked



Flexibility to make human decisions where the matching criteria is not smart enough



Ability to extend the functionality for more semantically driven decisions (handing off to a more sophisticated matching algorithm)



Capability to add code lists or other dimensions, population or respondent to finesse the decisions



Capture matching decisions that could be added to a Quality Statement for transparency



# Thank you

**Bill Norley**

CLOSER Programmer

**Hayley Mills**

Senior Metadata Officer



@CLOSER\_UK



closer@ucl.ac.uk

Explore the content of UK  
longitudinal studies

**discovery.closer.ac.uk**

