

Question driven harmonisation of data The variable cascade in practice

Jon Johnson & Bill Norley CLOSER, UCL Social Research Institute

December 2020

License: <u>CC BY 4.0</u>

2020



Partners

Hertfordshire Cohort Study

MRC National Survey of Health and Development

1958 National Child Development Study

1970 British Cohort Study

ONS Longitudinal Study

Whitehall II

Understanding Society: The UK Household Longitudinal study

Avon Longitudinal Study of Parents and Children

Southampton Women's Survey

Millennium Cohort Study

English Longitudinal Study of Aging

Health and Employment After Fifty







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

← → ♂ ☆	🖸 🗋 localhost:502	62/Equivalence/Results			♡☆ II\ 🗉 📽 =
Select Variable Equival	lences from Study	Display Items	Select Question Equivalences from Sweeps	Select Variable Equivalences from Sweeps	Question Matching for Sweeps in all Studies
Expected Outp	out				
Index - Allitems = System.Colle	× +				– ø ×
$\leftarrow \rightarrow \times \ \textcircled{a}$	🛛 🗋 localhost:502	62/Equivalence/Index			⊽☆ IN 🖸 📽 Ξ
Select Question Equiva	lences from Sweeps	Display Items	Select Question Equivalences from Sweeps	Select Variable Equivalences from Sweeps	Question Matching for Sweeps in all Studies
[{"text":"COVID-19 Wave 1","id":" [{"text":"COVID-19 Wave 1","id":"	'uk.iser 3f603afc-c53a-4609-9b9 'uk.iser 3f603afc-c53a-4609-9b94	4-ec238b2b8e8a","parent":"uk.iser"},{"text":"COVID-19 Wave 2 4-ec238b2b8e8a","parent":"uk.iser"},{"text":"COVID-19 Wave 2	","id":"uk.iser 9b054291-6aa6-4cd4-b4ad-1863b8a1aa0f","parent" ","id":"uk.iser 9b054291-6aa6-4cd4-b4ad-1863b8a1aa0f","parent"	":"uk.iser"},{"text":"COVID-19 Wave 3","id":"uk.iser 3b0f3340-5ce ":"uk.iser"},{"text":"COVID-19 Wave 3","id":"uk.iser 3b0f3340-5ce	9-4f48-b56f-4f5fb9556c3e","parent":"uk.iser"}] 9-4f48-b56f-4f5fb9556c3e","parent":"uk.iser"}]
🔲 🛴 1970 British Cohort	Study			Index - AllItems = 1648	
Kon Longitudinal S	rt Study of Parents and Children			There are a number of methods will only provide a match where	for matching equivalences in Studies. The methods below there is 100% correlation
; 🔲 📗 Millennium Cohort S	Study			□ Match Text	
National Child Deve	elopment Study Health and Development			□ Match All Words	
National Survey of P	realth and Development			O Match Name Match Text	
s Southampton Wome	en's Survey			Please select the equivalences to equivalences	be analysed and then sweep(s) to be used with these
þ 🔳 🕼 Understanding Socie	ety				
i' 🛄 🥡 Wirral Child Health	and Development Study			Select a me of Equivalences to Opload	
				Browse No file selected.	Upload Equivalences
				or Add word(s) to search. You can add	or delete word(s) to search by using the appropriate buttons.
					Save
				currently living partner	Delete
				childcare home schooling	Delete
				currently pregnant	Delete
				vouchers meals schools	Delete
				work at home	Delete
				unhappy depressed	Delete

-)→ C' û	🛛 🗋 localhost:50	0262/Equivalence/Results				… ⊠ ☆	III\ 🗉 📽
Select Variable Ed	quivalences from Study	Display Items	Sweeps	Select Variable Equivalences from Sweeps	Question Matching for Swe	eps in all Studies	
Expected Ou	utput						
ItemType - My ASP.NET Ap	oplic: × +						- 0
-)→ ሮ ₪	🛛 🗋 localhost:50	0262/Equivalence/Index?fileName=sample text equivalences				▽ ☆	III\ 🗉 🔹
Select Variable Ec	quivalences from Study	Display Items	Select Question Equivalences from S	Sweeps	Select Variable Equivalences from Sweeps	Question Matching for Sweep	os in all Studies
("text":"COVID-19 Wave 1","ii ("text":"COVID-19 Wave 1","ii	d":"uk.iser 3f603afc-c53a-4609-9b94-ec	AIIILETIIS IO40 :238b2b8e8a", "parent": "uk.iser" },{"text": "COVID-19 Wave 2", "id": "uk.ise :238b2b8e8a", "parent": "uk.iser" },{"text": "COVID-19 Wave 2", "id": "uk.ise	r 9b054291-6aa6-4cd4-b4ad-1863b8a1aa0f", "parent": r 9b054291-6aa6-4cd4-b4ad-1863b8a1aa0f", "parent":	"uk.iser"},{"text":"CO\ "uk.iser"},{"text":"CO\	/ID-19 Wave 3","id":"uk.iser 3b0f3340-5ce9-4f48-b56f-4f5fb9556c3e", /ID-19 Wave 3","id":"uk.iser 3b0f3340-5ce9-4f48-b56f-4f5fb9556c3e",	,"parent":"uk.iser"}] ,"parent":"uk.iser"}]	
ID Search Term	Question Name	Question Text	Variable Name		Variable Text	Agency	Study
1 currently living	✓ qi_couple	Are you currently living with a partner?		ca_couple	Living with a partner	uk.iser	COVID-19 Wave 1
	✓ qi_couplewsh	Are you currently living with a partner?		cb_couplewsh	Living with a partner	uk.iser	COVID-19 Wave 2
	🗹 qi_couplewsh	Are you currently living with a partner?		cc_couplewsh	Living with a partner	uk.iser	COVID-19 Wave 3
D Search Term	Question Name	Question Text	L	<u></u>	Variable Text	Agency	Stud,
childcare home schooling	✓ 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 schoolir	ng uk.iser	COVID-19 Wave 1
y	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 schoolir	ng uk.iser	COVID-19 Wave 2
	v qi_hrchsch	Have you reduced the hours you spend in your paid jo home schooling?	b because of the time you spend doing childcare or	cb_hrchsch	Employment hours change due to homes	chooling uk.iser	COVID-19 Wave 2
	✓ qi_workchsch	Have you adapted the work schedule of your paid job home schooling?	because of the time you spend doing childcare or	cb_workchsch	Work pattern change due to homeschooli	ing uk.iser	COVID-19 Wave 2
	qi_workchsch	Have you adapted the work schedule of your paid job home schooling?	because of the time you spend doing childcare or	cc_workchsch	Work pattern change due to homeschooli	ing uk.iser	COVID-19 Wave 3
	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 schoolir	ng uk.iser	COVID-19 Wave 3
	✓ qi_hrchsch	Have you reduced the hours you spend in your paid jo home schooling?	b because of the time you spend doing childcare or	cc_hrchsch	Employment hours change due to homes	chooling uk.iser	COVID-19 Wave 3
D Search Term	Question Name	Question Text		Variable Name	Variable Text	Agency	Study
currently pregnant	✓ qi_preg	Are you currently pregnant?		ca_pregnow	Whether pregnant	uk.iser	COVID-19 Wave 1
	✓ qi_pregnow	Are you currently pregnant?		cb_pregnow	Whether pregnant	uk.iser	COVID-19 Wave 2
	✓ qi_pregnow	Are you currently pregnant?		cc_pregnow	Whether pregnant	uk.iser	COVID-19 Wave 3

Variable Name

ca_blhhearn_period

Variable Text

Current household earnings period

Current household earnings amount

Baseline: household earnings period

Study

COVID-19 Wave 1

COVID-19 Wave 1

COVID-19 Wave 1

Agency

uk.iser

uk.iser

uk.iser

ID Search Term

work at home

4

Question Name

✓ qi_hhearn_period

✓ qi_hhearn_amount

✓ qi_blhhearn_period

Question Text

tell us an approximate amount. Per

tell us an approximate amount. ... Pounds

you are not sure, please tell us an approximate amount. Per

Thinking about everyone living with you at the moment, what is the total take-home pay/earnings of your

Thinking about everyone living with you at the moment, what is the total take-home pay/earnings of your

Thinking about everyone who was living with you in January/February 2020, what was the usual total takehome 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

household now? Please only include earnings from paid work or self-employment. If you are not sure, pleaseca_hhearn_period

household now? Please only include earnings from paid work or self-employment. If you are not sure, please ca_hhearn_amount

€)→	C 🗅		localhost:50	262/Equivalence/Results				⊍	☆	\ ⊡ 📽 ≡
Select Variable Equivalences from Study Display Items				Display Items	Selec	Select Question Equivalences from Sweeps Select Variable Equivalences from Sweeps		Question Matching for Sweeps in all S		os in all Studies
Exp	ected C	Dutput								
ltem]	ype - My ASP.NET	Applic × +								– o ×
€→	C 🕜		localhost:50)262/Equivalence/Display?fileName=sample text equi	valences			⊌	☆	II\ ⊡ 📽 ≡
	Select Variable	e Equivalences from Stu	ıdy	Display Items	Selec	t Question Equivalences from Sweeps	Select Variable Equivalences from Sweeps	Question Ma	tching for Sweep	s in all Studies
Iden	tifying Ec	uivalent - All	ltems =	1648						
[{"text":	"COVID-19 Wave 1	","id":"uk.iser 3f603afc-c53	a-4609-9b94-ec2	238b2b8e8a", "parent":"uk.iser"},{"text":"COVID-19 Wave 2","id	":"uk.iser 9b054291-6aa6-4	cd4-b4ad-1863b8a1aa0f", "parent":"uk.iser"},{"text":"COVI	ID-19 Wave 3", "id":"uk.iser 3b0f3340-5ce9-4f48-b56f-4f5fb9556c3e", "p	arent":"uk.iser"}]		
ID	Search Term	Name	Text		Variable Name	Variable Text	Variable Item	Que	stion Name	Question Text
1	currently living	ca couple	Living with a	a partner	ca couple	Living with a partner	uk.iser.ukhls:7113e2c5-64b6-4f36-bf49-2f5f88	27359:3 gi c	ouple	Are you currently living
	partner				cb_couplewsh	Living with a partner Living with a partner	uk.iser.ukhls:12d52199-3699-4286-ad34-3ba17 uk.iser.ukhls:cdc9ea7c-d0ea-418f-879f-c2175(5665c73:3 qi_c 18f906d:3 qi c	ouplewsh	Are you currently living Are you currently living
ID	Search Term	Name	Text		Variable Name	Variable Text	Variable Item	Que	stion Name	Question Text
2	childcare home	ca timechcare	Time spent	on childcare or home schooling	ca timechcare	Time spent on childcare or home schooling	uk iser.ukbls:6e684650-c3c2-4735-a99f-71e20c	Bed3ef:3 ai ti	mechcare	About how many hour
-	schooling			on enhance of none schooling	ch timecheara	Time spent on childcare or home schooling		1772-072.2 ai ti	mocheara	About how many hour
					cb_hrchsch	Employment hours change due to homeschooling	uk.iser.ukhls:a16d4034-a103-4987-9bd5-8e561	7134eac:3 qi_h	rchsch	Have you reduced the F schooling?
					cb_workchsch	Work pattern change due to homeschooling	uk.iser.ukhls:16e88ee2-d564-4cb3-9e7b-74a8e	44577be:3 qi_w	orkchsch	Have you adapted the v schooling?
					cc_workchsch	Work pattern change due to homeschooling	uk.iser.ukhls:6059e6b6-1bd3-41f6-b70f-7c44a	d8d461b:3 qi_w	orkchsch	Have you adapted the \ schooling?
					cc_timechcare	Time spent on childcare or home schooling	uk.iser.ukhls:1caa1281-fadc-40c3-a62c-46fc06	38a441:3 qi_ti	mechcare	About how many hour:
					cc_hrchsch	Employment hours change due to homeschooling	uk.iser.ukhls:64cec6cb-1aff-4629-8b60-e0bd0	b75a482:3 qi_h	rchsch	schooling?
ID	Search Term	Name	Text		Variable Name	Variable Text	Variable Item	Que	stion Name	Question Text
3	currently pregnar	nt ca_pregnow	Whether pre	egnant	ca_pregnow	Whether pregnant	uk.iser.ukhls:5e311454-ae63-47e6-b0e6-a6410	689bb4f:3 qi_p	reg	Are you currently pregr
					cb_pregnow	Whether pregnant	uk.iser.ukhls:8b191d72-76f0-412b-b4f6-aed44	c7d578c:3 qi_p	regnow	Are you currently pregr
					cc_pregnow	Whether pregnant	uk.iser.ukhls:ba498ba2-bbec-43e4-8a52-4b0b4	189a09e3:3 qi_p	regnow	Are you currently pregr
ID	Search Term	Name	Text		Variable Name	Variable Text	Variable Item	Que	stion Name	Question Text
4	work at home	ca_hhearn_period	Current hou	isehold earnings period	ca_hhearn_period	Current household earnings period	uk.iser.ukhls:78206333-a4f4-4cfa-a32d-34bfd1	d3f326:3 qi_h	hearn_period	I hinking about everyor household now? Please
					ca_hhearn_amount	Current household earnings amount	uk.iser.ukhls:cb52b378-0625-4be1-84c9-00bb7	7d04df2e:2 qi_h	hearn_amount	Thinking about everyor household now? Please us an approximate amc Thinking about everyor
					ca_blhhearn_period	Baseline: household earnings period	uk.iser.ukhls:4cf6b8b9-38b3-4600-bc37-f2e53	≥261b13:3 qi_b	hhearn_period	pay/earnings of your h employment, after tax,
					ca_wah	Working at home	uk.iser.ukhls:cfe01df0-2f77-4f26-8ad9-ae0cc0	e745ea:3 qi_w	ah	please tell us an approx During the last four we
					ca hstime childe	Time taken - child e	uk.iser.ukhls:aa31f01a-0764-4cca-8629-08d149	93a40bd:3 aih	stime	I hinking about the situ

$\left(\leftarrow ight) ightarrow$ C $ ightarrow$	🛛 🗋 localhost:5	O □ localhost:50262/Equivalence/Results O ☆							
Select Variable Equivalence	es from Study	Display Items	Select Question Equivalences from Sweeps	Select Variable Equivalences from Sweeps	Question Matching for Sweeps in all Stu	dies			
Expected Outpu	t								

Image: International Select Variable Equivalences from Study Select Variable Equivalences from Study Display Items Select Variable Equivalences from Study Display Items Select Variable Equivalences from Study Display Items Select Variable Equivalences from Sweeps Question Matching for Sweeps in all Studies

Expected Output

[{"text":"COVID-	{"text":"COVID-19 Wave 1","id":"uk.iser 3f603afc-c53a-4609-9b94-ec238b2b8e8a","parent":"uk.iser"},{"text":"COVID-19 Wave 2","id":"uk.iser"},{"text":"COVID-19 Wave 3","id":"uk.iser 3b0f3340-5ce9-4f48-b56f-4f5fb9556c3e","parent":"uk.iser"}]								
Uniqueld	Name currently living partner	Description Living with a partner		Topic Family and social networks (CV19)		COVID-19 Wave 1 ca_couple	COVID-19 Wave 2 cb_couplewsh	COVID-19 Wave 3 cc_couplewsh	ĺ
Uniqueld 2	Name childcare home schooling	Description Work pattern change due to homeschooling		Topic Employment and income (CV19) Family and social networks (CV19) Employment and income (CV19)		COVID-19 Wave 1 ca_timechcare	COVID-19 Wave 2 cb_timechcare cb_hrchsch cb_workchsch	COVID-19 Wave 3 cc_workchsch cc_timechcare cc_hrchsch	
Uniqueld	Name	Description		Iopic Pregnancy (CV19)		COVID-19 Wave 1	COVID-19 Wave 2 ch. pregnow	COVID-19 Wave 3	
Uniqueld 4	Name work at home	Description Current household income amount		Topic Employment and income (CV19) Employment and income (CV19) Education (CV19)		COVID-19 Wave 1 ca_hheam_period ca_hheam_period ca_bheam_amount ca_blheam_period ca_wah ca_hstime_childe ca_hstime_childb ca_hstime_childb ca_hstime_childf ca_hstime_childd ca_hstime_childd ca_hstime_childd ca_hstime_childd ca_hstime_childd ca_hstime_childd ca_hstime_childb ca_stime_childb ca_stime_childb ca_stoolwork_childb ca_schoolwork_childb ca_sc	COVID-19 Wave 2 cb_hhearn_amount cb_blhhearn_period cb_hhincome_amount cb_deskspace cb_blhhearn_period cb_hhincome_period cb_hhincome_period cb_blwah cb_wah	COVID-19 Wave 3 cc_hhincome_amount cc_prodch cc_blhearn_period cc_hhearn_period cc_wah cc_hhincome_period cc_hhearn_amount cc_blwah cc_blhhearn_amount	





⊕- □ Demographics		?			
Housing and local environment					
			COVID-19 Wave		
			1	COVID-19 Wave 2	COVID-19 Wave 3
Lealth care			ca indresp w	cb indresp w	cc indresp w
Family and social networks	+ Ne	w health conditions		cb_hcondnew_cv22	cc_hcondnew_cv22
	dia	gnosed - An emotional,			
E Employment and income	ner	rvous or psychiatric			
Expectations, attitudes and beliefs	pro	blem			
+- Child development		w health conditions		ch beondnew cy5	cc bcondnew cv5
+- 🗅 Life events	dia	anosed - Angina		cb_nconditew_cvo	
⊕- ⊡ Omics	did	griobou / ligiliu			
🕀 🗅 Pregnancy	+ Ne	w health conditions		cb_hcondnew_cv12	cc_hcondnew_cv12
- 🗅 Administration	dia	gnosed - Any kind of liver			
E- COVID-19	COI	ndition			
— 🗅 Demographics (CV19)	- No	w bealth conditions		ch beendnew cy27	cc bcondnew cv27
 — D Housing and local environment (CV19) 		aposed - Are very		cb_nconditew_cv27	cc_ncondnew_cv27
Physical health (CV19)		griosed - Ale very			
— 🗀 Mental health and mental processes (CV'	000	Siweight			
— 🗀 Health care (CV19)	+ Ne	w health conditions		cb_hcondnew_cv2	cc_hcondnew_cv2
— 🗀 Health behaviour (CV19)	dia	gnosed - Arthritis			
 Eamily and social networks (CV19) 		w boolth conditions		ah haandnaw avi	as beendrow out
— 🗀 Education (CV19)		anosod - Asthma		cb_ncondnew_cv1	cc_ncondnew_cv1
— 🗀 Employment and income (CV19)	ula	griosed - Astrina			
 — D Expectation, attitudes and beliefs (CV19) 	+ Ne	w health conditions		cb_hcondnew_cv13	cc_hcondnew_cv13
- Child development (CV19)	dia	gnosed - Cancer or			





Demographics Housing and local environment		<u>.m</u> ?		
		B 11 1 1 1 1 1		
	+	Participant had loss of sense	covid1_2086	covid2_1008
		of smell or taste	covid1_2084	
Health care			covid1_2082	
+- D Health behaviour			covid1_2087	
🕂 🗀 Family and social networks			covid1_2080	
🕂 🗀 Education			covid1_2083	
🕂 🗀 Employment and income			covid1_2088	
🕂 🗀 Expectations, attitudes and beliefs			covid1_2081	
+- 🗀 Child development			covid1_2085	
+- 🗀 Life events	+	Participant had nausea and/or	covid1 2016	covid2 1001
		vomiting	covid1_2018	001142_1001
			covid1 2015	
			covid1 2017	
			covid1 2010	
			covid1 2011	
Demographics (CV19)			covid1 2013	
 Housing and local environment (CV19) 			covid1 2014	
Physical health (CV19)			covid1 2012	
— D Mental health and mental processes (CV'			_	
— 🗀 Health care (CV19)	+	Participant had numbness or	covid1_2216	covid2_1021
— 🗀 Health behaviour (CV19)		tingling	covid1_2211	
— 🗀 Family and social networks (CV19)			covid1_2214	
- C Education (CV19)			covid1_2212	
 Employment and income (CV19) 			covid1_2213	
 Expectation, attitudes and beliefs (CV19) 			covid1_2218	
			covid1_2210	

Summary

 \checkmark

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





Bill Norley CLOSER Programmer

Hayley Mills Senior Metadata Officer



Explore the content of UK longitudinal studies

discovery.closer.ac.uk

