

the new Lifelines dataplatform

T.R. de Jong¹, G.L. Faber¹, R. van Vliet², M.A. Swertz³, A. Dotinga¹

¹Lifelines Cohort & Biobank, Roden, the Netherlands, www.lifelines.nl | research@lifelines.nl

²Trivento BV, Amersfoort, the Netherlands, www.trivento.nl | blij@trivento.nl

³Molgenis GCC, Groningen, the Netherlands, www.molgenis.org | molgenis-support@umcg.nl



OUR MISSION

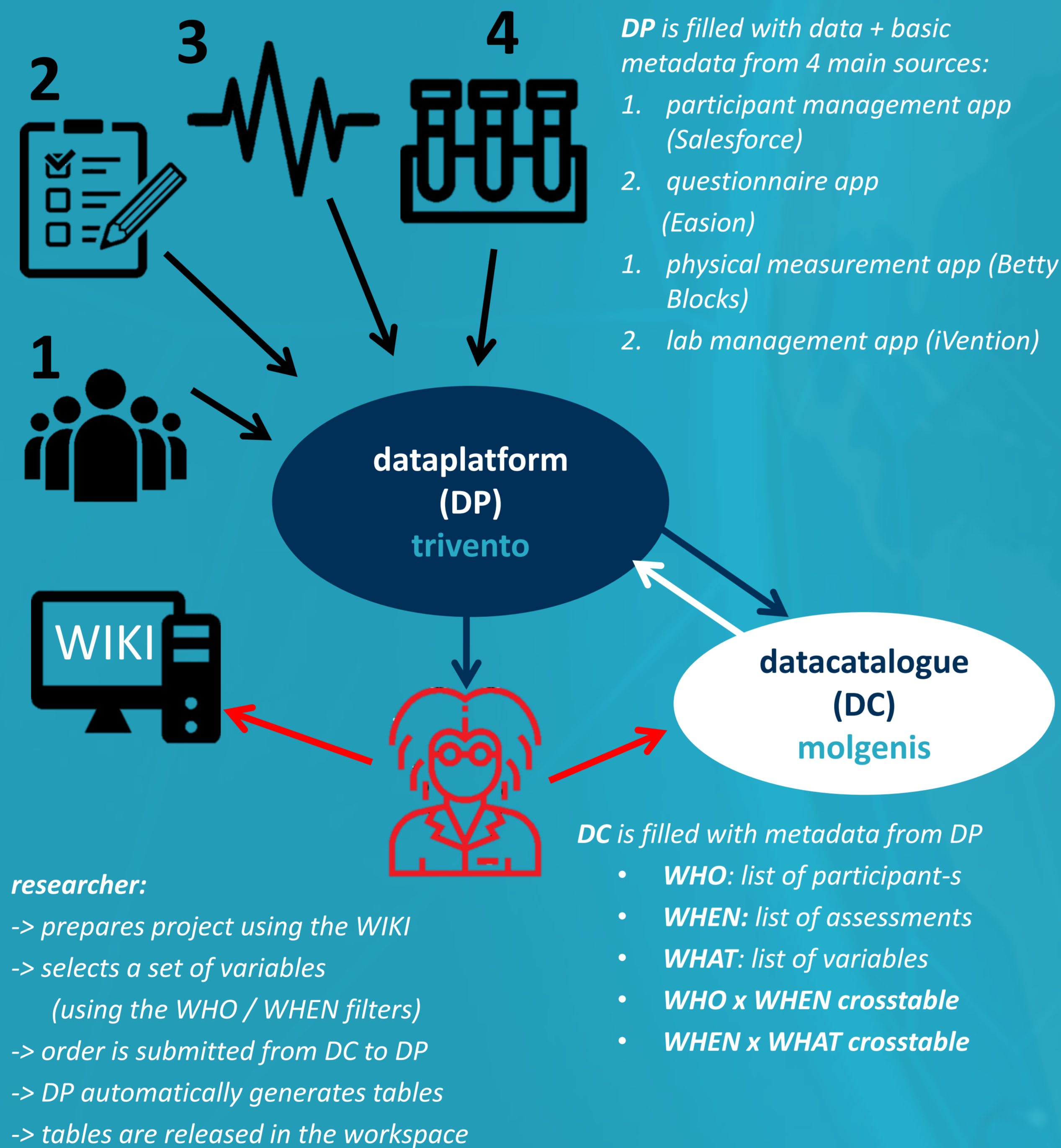
- Lifelines Cohort Study and Biobank collects longitudinal **DATA** and **SAMPLES** from >167,000 inhabitants of the northern Netherlands (including children and elderly)
- our mission is to make our rapidly expanding collection available for all researchers working in the field of **HEALTHY AGEING**
- our **OPEN PROTOCOL** allows researchers to collect additional data and samples from participants, to gain knowledge of healthy ageing

LINKS

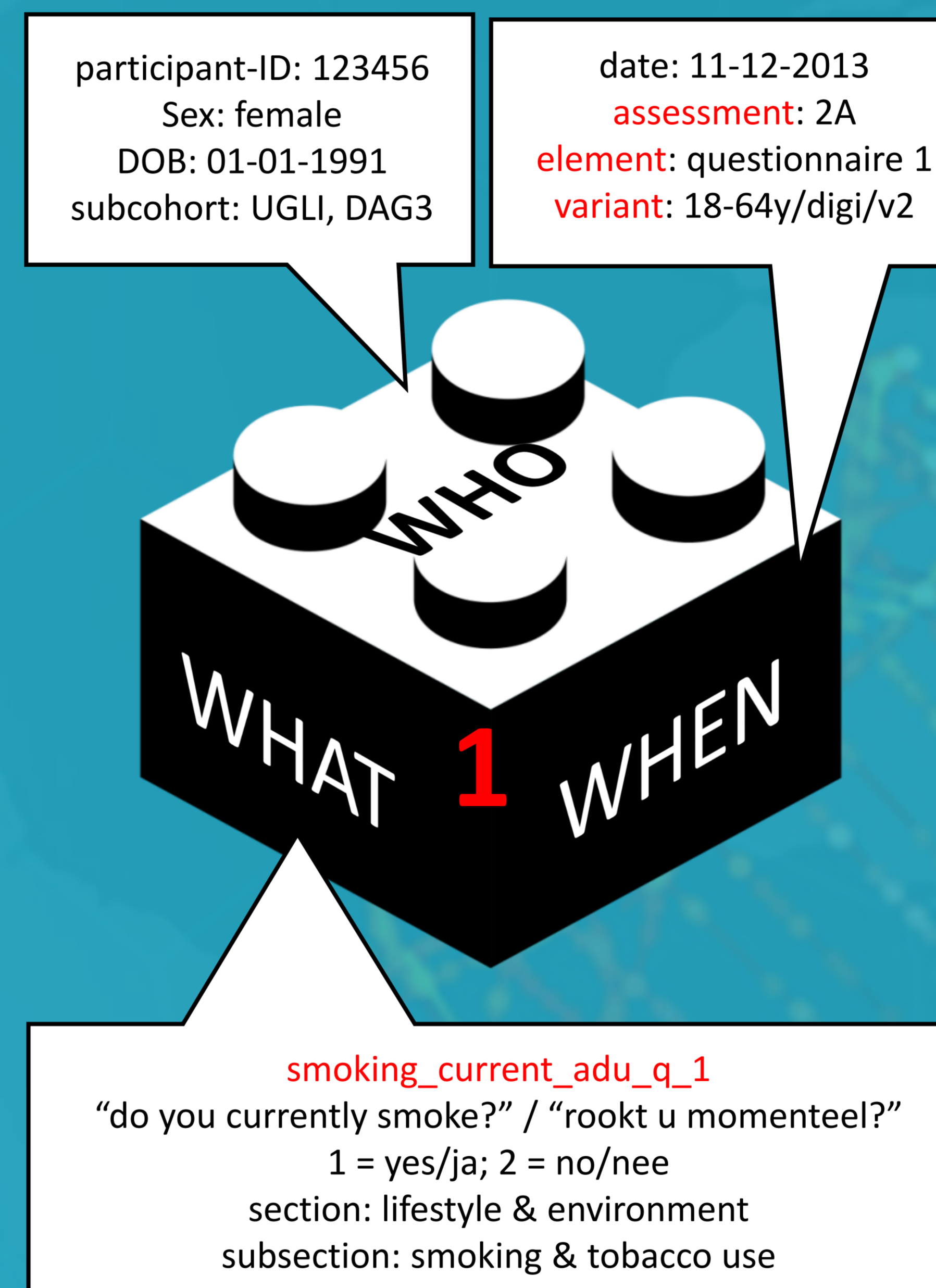
<https://data-catalogue.lifelines.nl>

<http://wiki-lifelines.web.rug.nl>

OUR INFRASTRUCTURE



OUR DATA MODEL



THE “WHEN” DIMENSION

ASSESSMENT:

data/sample collection (project) with a common goal

ELEMENT:

part of an assessment taking place on 1 day/date

- > 1 visit
- > 1 questionnaire

VARIANT:

part of an element with a highly standardized protocol

- > 1 set of selection criteria
- > 1 set of variables
- > 1 method

assessment a

element 1	element 2
variant 1	variant 1
variant 2	variant 2
variant 3	variant 3
variant 4	variant 4

assessment b

element 1	element 2
variant 1	variant 1
variant 2	variant 2
variant 3	variant 3
variant 4	variant 4

(FAIR) AIM 1

to provide researchers in advance with a **COMPLETE OVERVIEW** of:

- our available participants/subcohorts
- our available assessments
- our available variables
- the number of participants per subcohort x assessment x variable
- our (rationale for certain) protocols

(FAIR) AIM 2

to practice **DATA MINIMALIZATION** by:

- building datasets that are customized on-demand
- helping researchers to filter out participants or subcohorts (by profile) and/or assessments
- allowing researchers to select individual variables (not fixed sets)

(FAIR) AIM 3

to **STRUCTURE** our datasets to facilitate analyses, including info about:

- missing vs. not collected data
- responders vs. non-responders
- “hidden” protocol changes
- basic (must-have) metadata for each participant & variable
- protocol description for each variant

(FAIR) AIM 4

to **RAPIDLY INCORPORATE** incoming data collections, including:

- regular data from our 4 main sources
- linked data from other registries (e.g. via postal codes)
- secondary (coded, sum-scored) data
- cleaned & qc'd datasets
- data from new sources, e.g. apps

SUPPORT

Lifelines is supported by:

UMC Groningen
Groningen University
Dutch Ministries of Health, Welfare & Sports / Economic Affairs
3 Northern Provinces:
Drenthe, Friesland and Groningen

Molgenis is supported by:

EU CAN-connect (H2020-824989)
EOSC-life (H2020-824087)