SND Svensk Nationell Datatjänst

University of Gothenburg
Karolinska Institutet
Lund University
Stockholm University
Swedish University of Agricultural Sciences
Umeå University
Uppsala University

DDI in 7 minutes

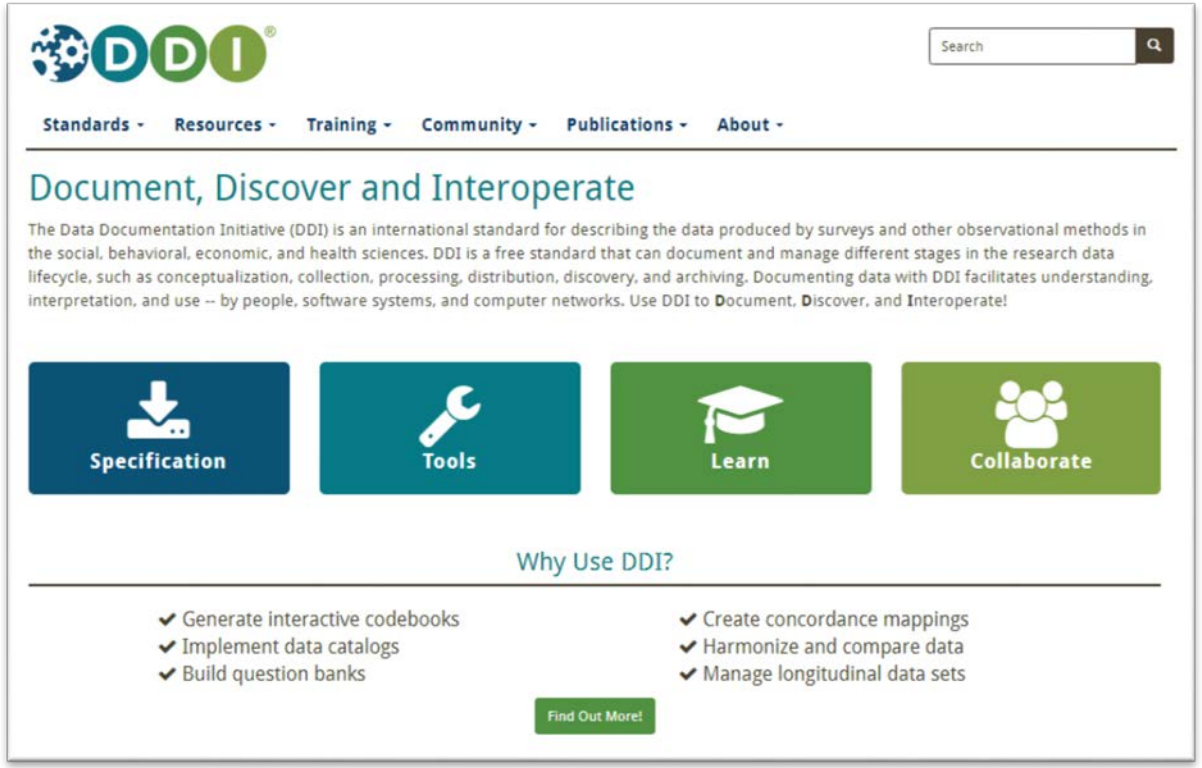
2018-10-08

Olof Olsson, IT Architect

GÖTEBORGS UNIVERSITET

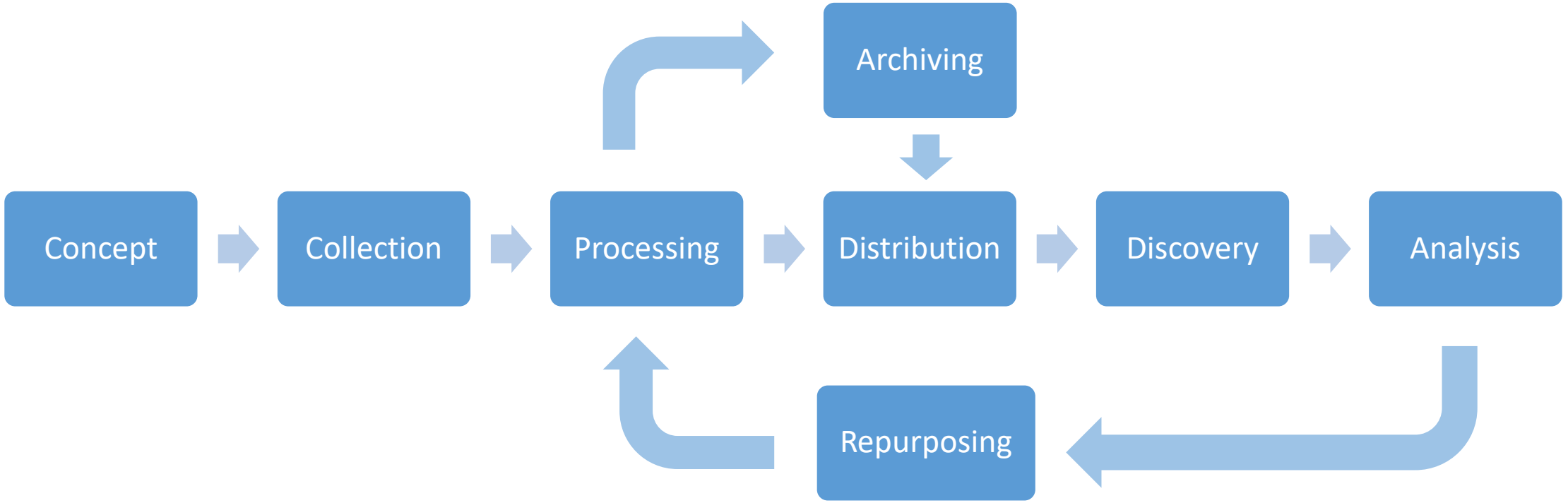
What is DDI?

- XML metadata standard for describing datasets
- Have grown to cover:
  - Data collection
  - Questionnaires
  - Geographical structures
  - Controlled vocabularies



https://www.ddialliance.org

DDI Lifecycle



Members and users of DDI

- Data archives
- Universities
- Statistical agencies
- Software companies

https://www.ddialliance.org/community

Controlled vocabularies

Some examples of DDI Controlled vocabularies:

Analysis Unit
Describes the entity being analyzed in the study or in the variable.

Data Type
Identifies the type of data, which has a bearing on the acceptable data values, the operations that can be performed with the data, and the ways in which the data are stored.

Mode of Collection
The procedure, technique, or mode of inquiry used to attain the data.

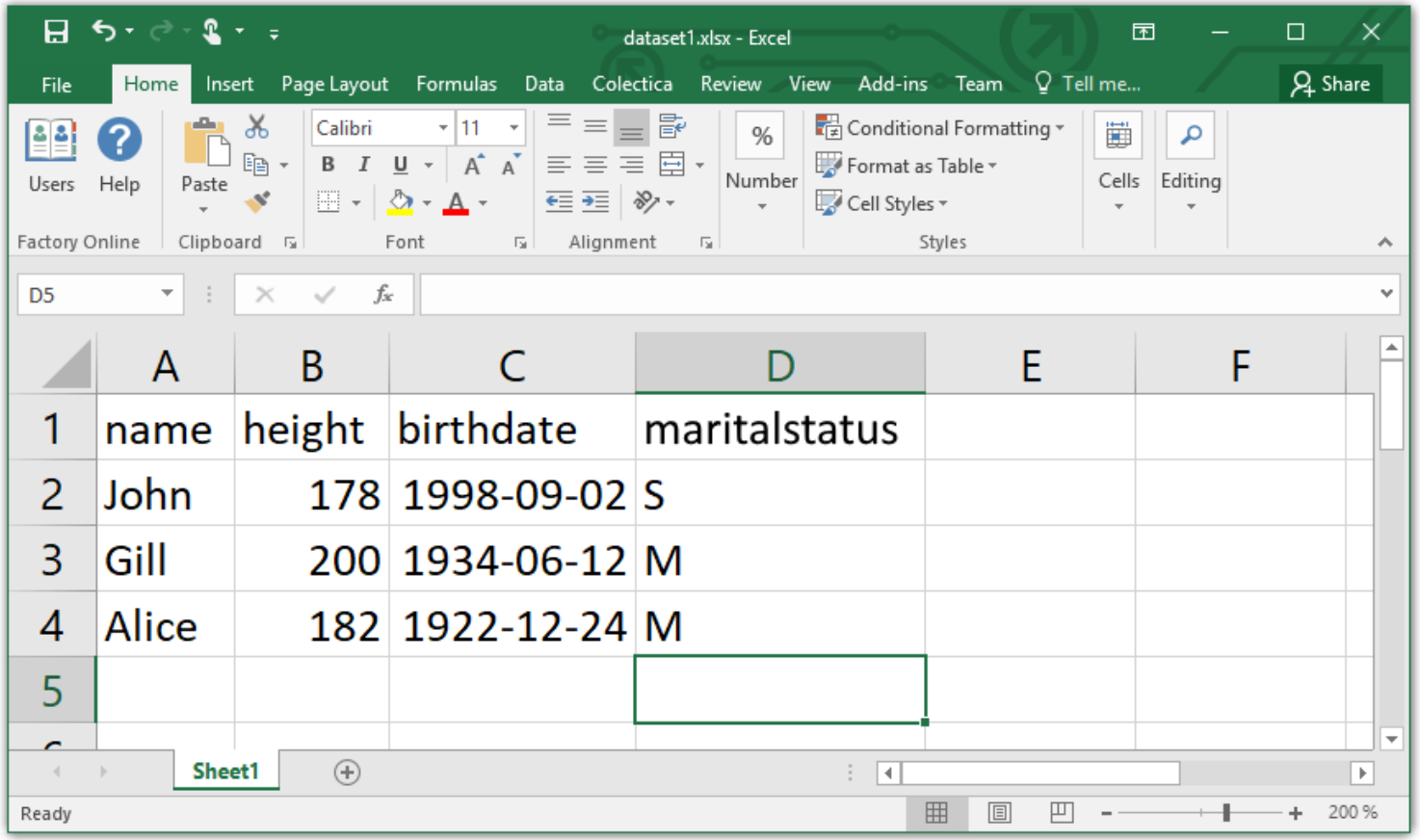
Sampling Procedure
Includes a typology of sampling methods.

Software Package
Indicates the statistical software package used in the production/processing/dissemination of the data.

Time Method
Describes the time dimension of the data collection.

Type of Instrument
Includes a typology of data collection instruments.

Data ⟷ Metadata



Example: dataset1

name	height	birthdate	maritalstatus
John	178	1998-09-02	S
Gill	200	1934-06-12	M
Alice	182	1922-12-24	M

Variable example

name	height	birthdate	maritalstatus
John	178	1998-09-02	S
Gill	200	1934-06-12	M
Alice	182	1922-12-24	M

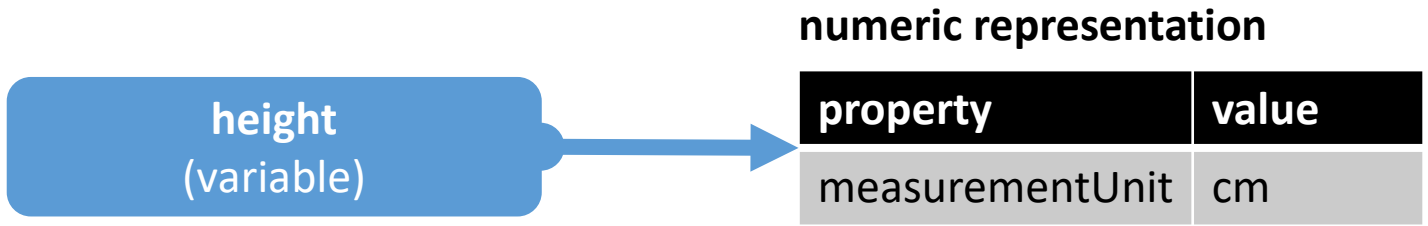
Different variable representation types

text	numeric	date	code
name	**height**	**birthdate**	**maritalstatus**
John	178	1998-09-02	S
Gill	200	1934-06-12	M
Alice	182	1922-12-24	M

Variable example

name	height	birthdate	maritalstatus
John	178	1998-09-02	S
Gill	200	1934-06-12	M
Alice	182	1922-12-24	M

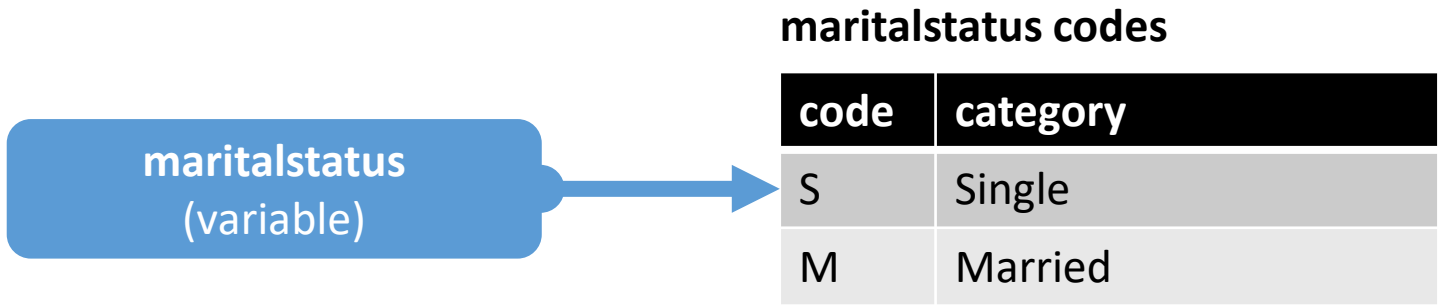
Variables with different unit of measure



Variable with code representation

name	height	birthdate	maritalstatus
John	178	1998-09-02	S
Gill	200	1934-06-12	M
Alice	182	1922-12-24	M

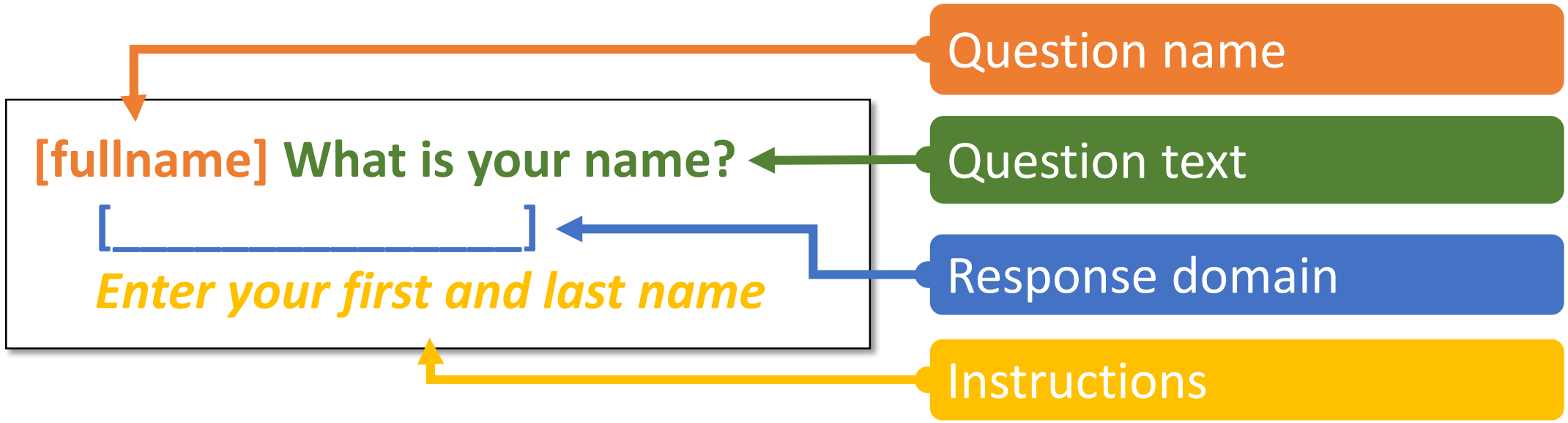
Variable - code representation



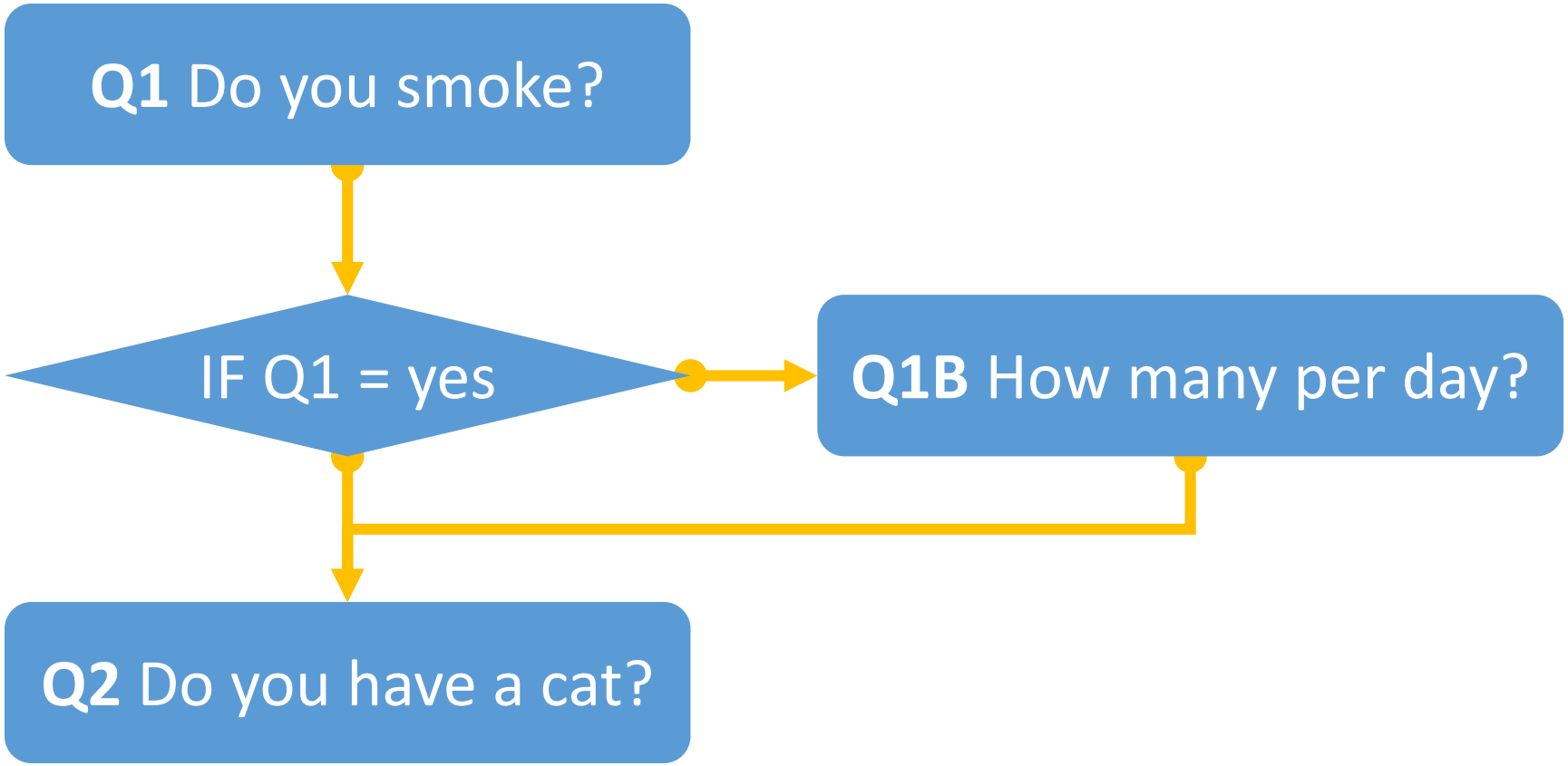
maritalstatus codes

code	category
S	Single
M	Married

Question item example



Questionnaire routing



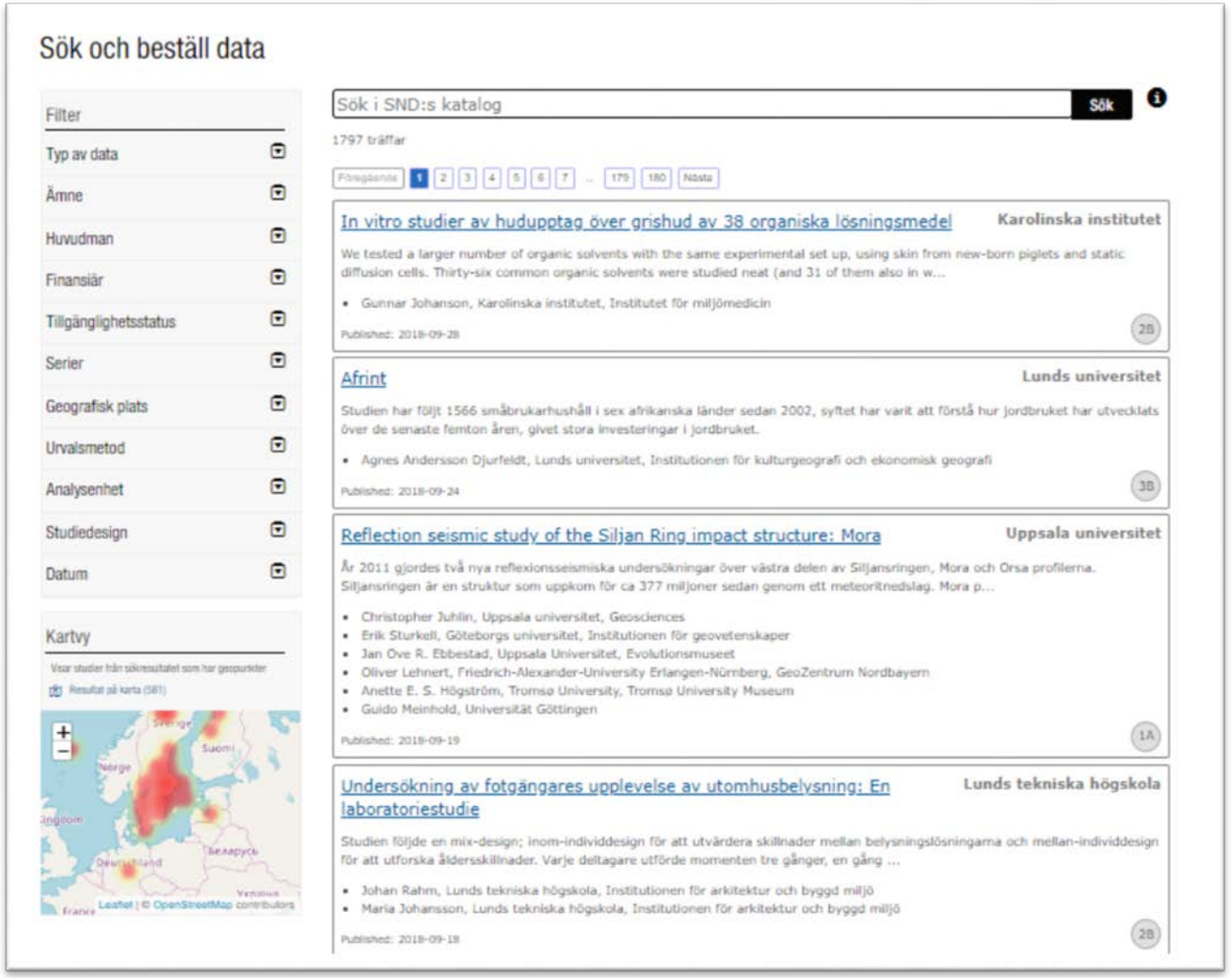
DDI at SND

Findability

- **Citation**: id, title, description, persons, organisations
- **Coverage**: time, place, subject
- **Collection**: method, population

Administrative metadata

- Funding
- Contacts
- Other identifiers



DDI overlap with other standards