Q 2016
European Conference on Quality in Official Statistics

Measuring occupations: respondent's self- identification from a large database

SPECIAL SESSION: Synergies for Europe's Research Infrastructures in the Social Sciences and Official Statistics (SERISS)

THU 2 June 2016

Kea Tijdens

University of Amsterdam, NL

k.g.tijdens@uva.nl

INE Instituto Nacional de Estadística eurostat

Madrid, May 31 - June 3

Setting the scene

The national stocks of job titles are …

- large >> 10,000's of job titles in any national labour force
- unstructured >> vague boundaries between job titles
- unlimited >> no fixed list, many entries and exits over time

The challenge

- to classify job titles into ISCO-08 classification of occupations
- & ... to do so consistently across countries

Occupational titles vs job titles

- job titles: within organisational context
- occupational titles: beyond organisational context

Q 2016
European Conference on Quality in Official Statistics

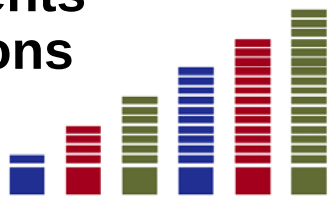
‘What is your occupation?’

Open-ended questions: textbox

- Textbox: predominantly used in surveys
- Office coding needed: expensive and time-consuming, though increasingly high quality coding software and auto-coders
- Coding problems: vague or highly aggregated titles

Closed questions: dictionaries

- Brief list (max 10 entries):
  predominantly used in postal surveys >> aggregation bias
- Showcard (max 50 entries):
  predominantly used in face-to-face surveys
- Dictionary (unlimited number of entries): respondents self-select their occupation from a list of occupations



Closed question: the example of WageIndicator web survey

WageIndicator websites

- In 2001: website with and wage content - started in the Netherlands
- Today: web portal with websites national websites in 89 countries, all in national language(s)
- 2015: 32 million of visitors, most through search engines

WageIndicator multilingual, continuous web survey

- All websites invite visitors to complete
  - a long salary survey in return to free information provided, with lottery incentive
  - or a mini-survey to get a salary indication
- Long survey N = 215 630 with valid ISCO; mini-survey N = 296 313 with valid ISCO (2015)

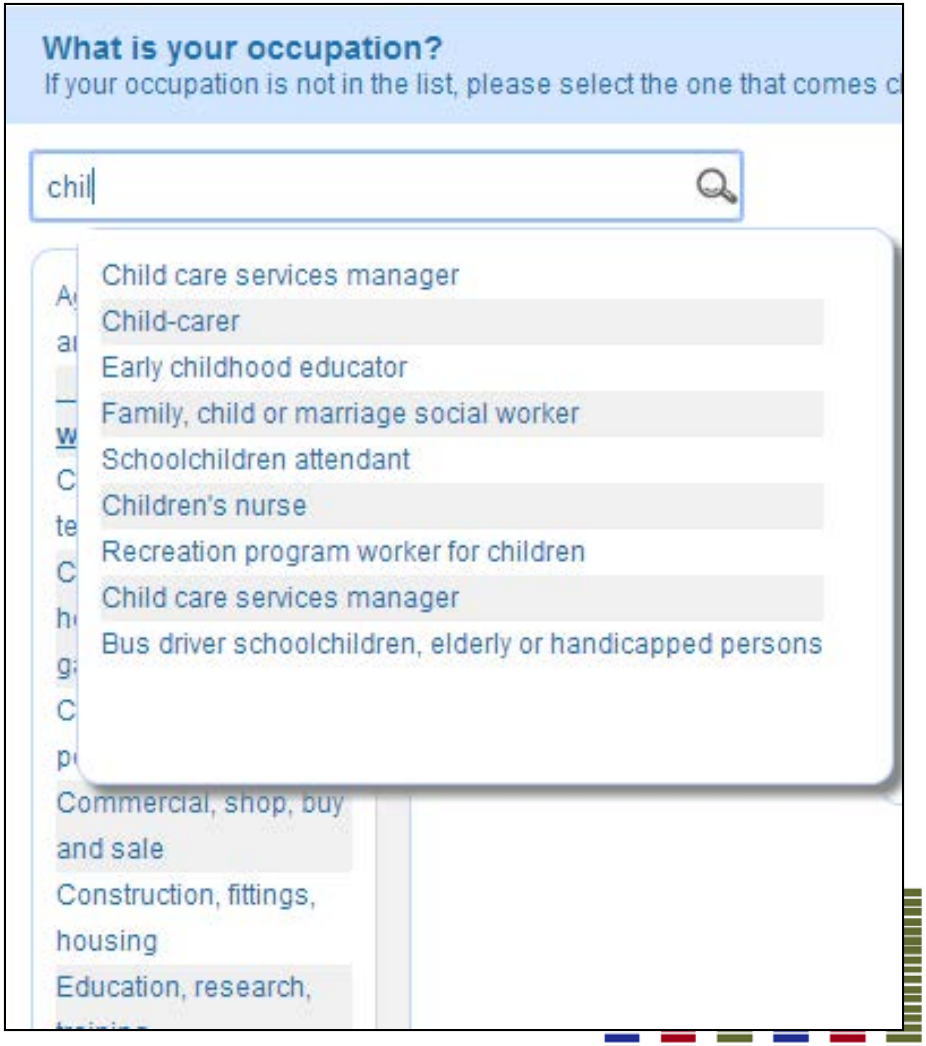
"What is your occupation?"

- Closed survey question (coding too expensive, particularly with more languages included)
- Respondents self-identify their occupation through
  - a search tree (IPod menu) and an autosuggest box (Google search type)
- Dictionary: a multi-lingual database of occupations, all coded 4 digit ISCO-08

History of this closed survey question

2001-'05	700 occupations in 1 language with 2-level search tree >> one page per level
2006-'09	1,100 occupations with 8 languages with 3-level search tree >> one page per level
2009-'15	1,600 occupations with 30 languages with 3-level tree on one page + autosuggest
20015-..	Database in 41 languages available on an API (Application Programming Interface)

Search tree (left), autosuggest (right)



Q 2016
European Conference on Quality in Official Statistics

SERISS project (2015-2019)

Extend the dictionary of occupations

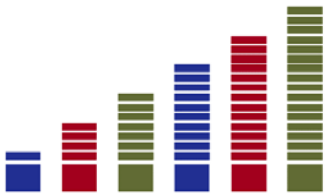
- To 99 countries with 47 languages
- To serve self-identication through search tree and through autosuggest
- Ensure that all occupational titles are well coded in ISCO-08

Make database available for survey holders

- Program an API for use in web surveys on desktop, tablet, smartphone
- Program an interface for use in CAPI surveys
- Make database downloadable in excel
- Availability: till end SERISS free of charge

Develop an occupation – industry prediction

- Depending on ticked occupation, a limited set of industries is shown for the survey question ‘In which industry do you work?’
- Aiming to reduce respondents’ time



Search tree vs autosuggest

Meta Data occupation API

- 23 – 30 May 2016
- 12 436 records from clicks and autosuggest >> 1990 respondents

Use of search tree versus autosuggest

- More than three in four uses the search tree (78%)
- More than one in five uses the text box (22%)
- Drop out rate (5%) [Note: drop out is common in this web survey]

Do not quote, because not controlled for mobile use

- Mobile users use only search tree

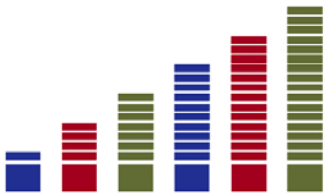
	complete	dropout	Total	%	% drop out
Search tree	1477	70	1547	78%	5%
Autosuggest	409	34	443	22%	8%
Total	1886	104	1990	100%	5%
% complete	95%	5%	100%		

Clicks in the search tree

Do respondents go back and forth in the search tree?

- **1547 respondents started the search tree**
- **54% found their occupation in three clicks**
- **14% went back and forth one time**
- **29% went back and forth more than one time**

back and forth	%
drop out	2.3
0	54.5
1	14.0
2	7.6
3	6.3
4	3.4
5	2.0
>5	9.8
total	99.9



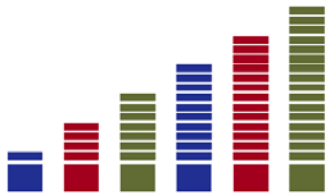
Q 2016
European Conference on Quality in Official Statistics

Response time

Response time in seconds

- After controlling for outliers (min 1 second, max 360 seconds) and for drop outs, response times of 1843 respondents were analysed
- Mean response time larger for autosugggest than for search tree (48 versus 44 seconds)
- Median response time larger for search tree than for autosugggest (26 versus 18 seconds)

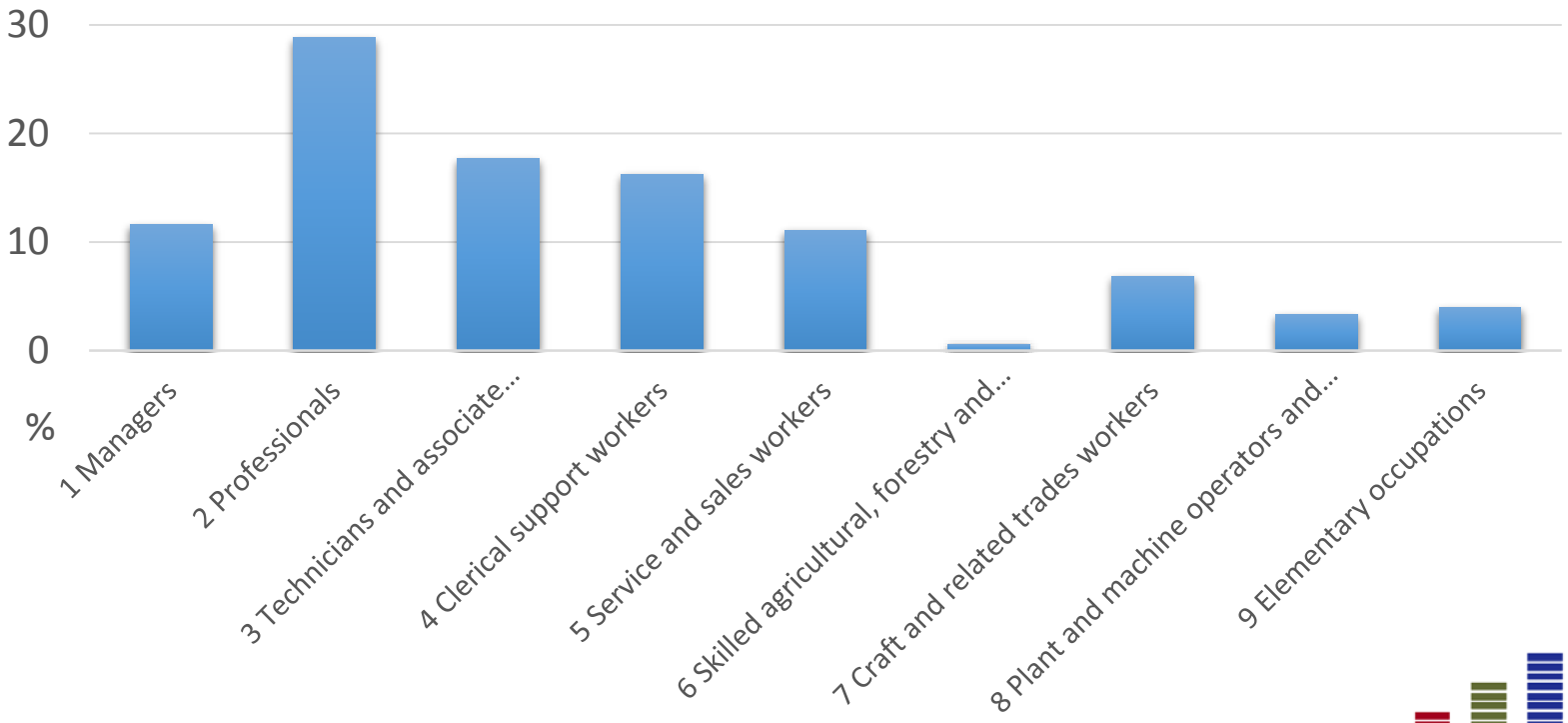
Seconds	Minimum	Median	Mean	Maximum
Search tree	1	26	44	352
Autosuggest	1	18	48	357



Selected occupations

Selected occupations ISCO-08

- The 1888 respondents selected 677 unique titles from the list of 1,600 titles
- Graph shows the distribution of sample for to 1-digit ISCO-08 classification



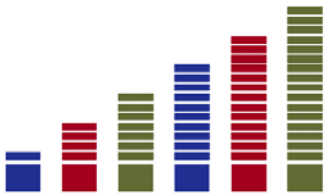
Q 2016
European Conference on Quality in Official Statistics

The end .....

Thank you for your attention ☺ ☺

Please try the demo:

http://tmt.centerdata.nl/jobcoder_demo/

Questions?

k.g.tijdens@uva.nl