



WageIndicator.org | Share and Compare Wages, Labour Law and Career

SSHOC

social sciences & humanities open cloud



The survey question measuring occupations solutions for multi-country web surveys

Kea Tijdens, research coordinator WageIndicator Foundation &
University of Amsterdam, Netherlands
& SSHOC WP3.2 Taskgroup leader 'Ontologies'

Presentation for GenPopWeb2 network, 12 March 2021



Brief intro: WageIndicator web survey

- I'm a sociologist, a retired research coordinator at University of Amsterdam and professor of Women's work in Erasmus University Rotterdam
- Since 2000 I'am also the scientific coordinator of WageIndicator Foundation
- WageIndicator is a web portal with websites in 196 countries in local language(s) with information concerning labour law, wages by occupation, minimum wages, career advice and alike, with 40 mln webvisitors p/y
- All national websites post a continuous web survey on work and wages >> multi-country survey in 47 languages
- The web survey has a question a question 'What is your job title?', using an occupation database for the answers
- The data is used to populate an online Salary Check, based on occupation



Measuring occupations: Why? How?

- **Why?**
 - **Occupation variable is used for research, e.g.**
 - Labour market >> wages, required skills, labour force composition
 - Identity >> occupation, social status
 - Occupational health and safety >> health risks per occupation
 - **Statistics Netherlands measures occupations for more than a century**
- **How?**
 - **Survey question >> “What is your job title”**
 - **Open answers versus closed answers**
 - **Open answers require office coding, using a coding index**
 - **Closed answers require a coded list of occupational/job titles, no office coding needed**

What do respondents do?

- **Respondents do know ...**
 - their job title from workplace, job evaluation, or job description
 - and they are mostly proud and eager to tell, and consistent across time
 - > very few 'don't know' and 'don't want to say' answers
- **Respondents do not know ...**
 - the ISCO-08 4 digit occupational units & how to classify their job title
 - what kind of answers the survey holder is looking for (some surveys solve this problem with instructions, e.g. school teacher)
 - -> crude answers -> aggregation heterogeneity (ISCO 1-4 dgt)
 - -> ambiguous, irrelevant answers, or abbrev. -> no coding
 - in CAWI & CATI interviewers can correct,
 - in PAPI and CAWI no correction, of these CAWI performs worst

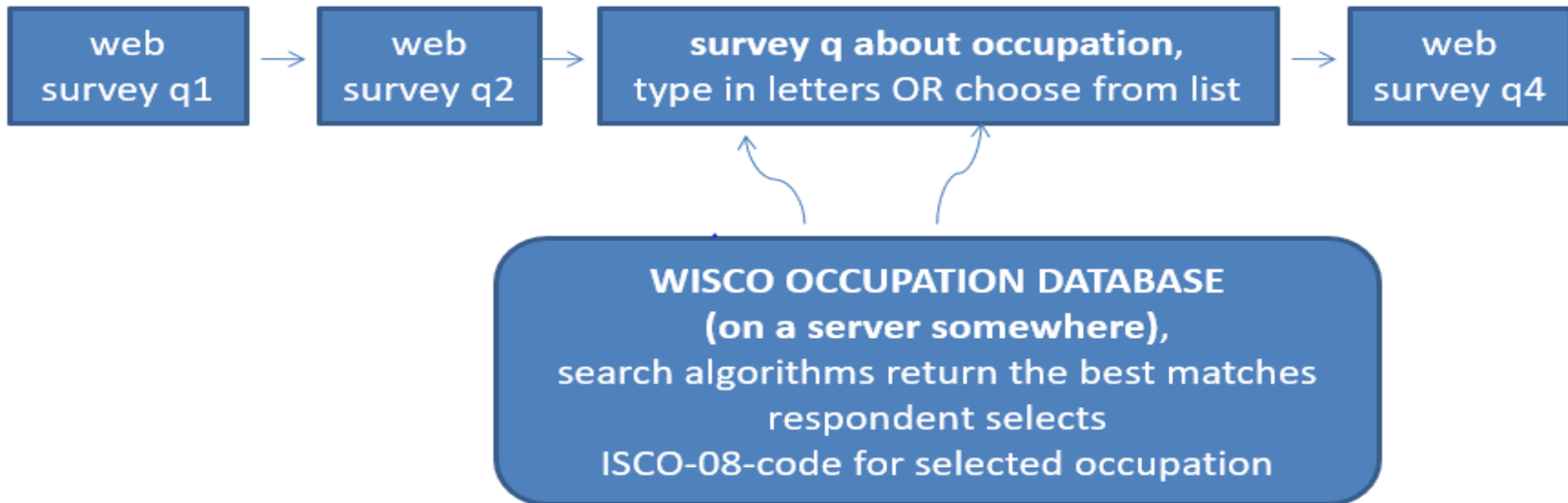


Can respondents self-identify?

- **Self-identification ...**
 - when using highly aggregated lists of occupations
 - > aggregation bias: respondents do not fit their job titles consistently into highly aggregated categories
 - when using disaggregated lists of occupations
 - > the smaller the distance to their own job title, the better they are able to classify their job title into an aggregated category, but difficult to search
- **... only in web surveys**
 - in WageIndicator web survey self identification of job title into a list of 1,800 occupational titles, all coded ISCO-08 4 dgt (called the WISCO database)
 - web surveys allow self-identification with disaggregated lists of occupations

Web survey

- **Application Programming Interface (API) (=Occupation database)**
 - Provided internet connection, any web survey can call upon the API





Measuring the long list of occupations ...

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, except licensed occ's
- unlimited >> no fixed list, many entries and exits over time
- distribution >> extremely skewed: many nurses, few C+ programmers

The classification challenge

- respondents report job titles, no occupational titles
- reported job titles have to be classified in an occupational classification
- specifically the long tail
- & ... to do so consistently across countries



Occupational classifications

- **National and international classifications**
 - 20th century: National Statistical Offices developed own classifications
 - 1958 - International Labour Organisation (ILO) developed International Standard Classification of Occupations (ISCO), updates 1968, 88, 08
 - Predominantly used by countries that had not an own classification
- **Harmonization in European Union**
 - 2009 - EU countries had to use ISCO-08 for Eurostat
 - 2012/13/14 - countries applied ISCO-08 for their Labour Force Surveys
 - ISCO-08 has become the international standard, but ...
- **ILO ...**
 - Hardly support staff, no discussion platform (poor compared to NACE class.)
 - ISCO-08 coding index is in English only



ISCO-08 classification logic

- **ISCO-08 is a 4-level hierarchical classification**
 - 10 major groups at the top of the hierarchy, based on skill levels (1-digit)
 - 436 occupational units at the bottom (4-digit)
 - ISCO-coding index has approx. 2,000 job titles (5-digit)
- **The challenge**
 - survey respondents report job titles (5-digit) & they do so reliably
 - >> coding 5-digit job titles into the 4-digit ISCO classification
 - >> to do so similarly across countries and languages
 - job titles are the same when the same tasks and duties are performed
 - no empirical tests of tasks, duties or required skill levels (beyond budget)
 - so ... coding in multi-country surveys is based on job title similarity only
 - requiring a multilingual coding index / dictionary



The measurement challenge – open q.

- **Office coding - problems**
 - coding problems for vague, aggregated, or company-specific titles
 - >> approx. 1-10% of responses is unidentifiable
 - >> approx. 1-10% has to be coded at higher aggregation level
 - office coding is expensive and time-consuming, though in few countries increasing high quality coding software and auto-coders
- **For multi-country surveys**
 - multi-country coding indexes do not exist
 - few examples of validating coding across countries
 - >> black box: are the same occupations coded similarly across countries ?



The measurement challenge – closed q.

- **Dictionaries with limited entries**
 - Brief list (max 10 entries): used in postal surveys >> aggregation bias
 - Showcard (max 50 entries): used in face-to-face surveys >> 'other' response
- **Dictionaries with many entries: Web surveys**
 - Dictionary (large number of entries): used in web surveys
>> respondents self-select their occupation from a list: Look-up databases
 - Self-identification by search tree (iPod menu)
or by autosuggest box (Google search type)
- **The challenge: how many entries?**
 - Measuring the long tail: will respondents identify 'synonyms' if the list does not include their job title?



Search tree <> text string matching

What is your occupation?
If your occupation is not in the list, please select the one that comes closest

- Agriculture, nature, animals, environment
- Child care, children, welfare, social work**
- Cars, mechanics, technicians, engineers
- Cleaning, housekeeping, garbage, waste
- Clerks, secretaries, post, telephone
- Commercial, shop, buy and sale
- Construction, fittings, housing
- Education, research, training
- Finance, banking, insurance

- Child care**
- Clergy
- Funeral service
- Maternity care
- Personal care
- Social work
- Support services (internal)
- Therapist
- counsellor, educator

- Au-pair
- Baby-sitter
- Child care services manager
- Child-carer
- Family day care worker
- Nanny
- Nursery assistant
- Nursery school teacher
- Out of school hours care worker

1 FIND YOUR OCCUPATION

- Web master, web manager
- Web designer**
- Web journalist
- Web programmer
- Web technician

1 FIND YOUR OCCUPATION

- Child care services manager
- Child care worker
- Children's nurse
- Early childhood educator**
- Family, child or marriage counsellor
- General practitioner for children (pediatrist)
- Recreation program worker for children
- Bus driver schoolchildren, elderly or handicapped persons
- Schoolchildren attendant



Can job titles be translated?

- **Job titles cannot be translated**
 - occupational titles cannot be translated beyond ISCO 4 digit
 - >> the black box continues
 - >> national coding indexes can be merged (provided ISCO-08 coding)
 - multi-country look-up database of coding indexes
- **Job titles can be translated**
 - job titles can be translated because similar job content, due to
 - globalisation of the economy -> need to understand occupations across countries
 - pressure towards cross-country standardization, e.g. QESH auditor
 - global equipment suppliers -> tasks in jobs become similar
 - multi-country look-up database of translated job titles

Database of merged coding indexes

- **Collecting coding indexes** (Tijdens & Kaandorp 2018)
 - NSOs in 99 countries: only 19 had a ISCO-08 with 5-digit titles
 - Austria 13,000+ occupational titles --- Finland 103 titles
 - we pooled the 19 indexes >> database with 70,489 titles
 - 9 indexes included non-existent ISCO-08 codes (10.3% of 70,489 titles)
- **Comparing the English translations**
 - using online dictionaries and Google Translate, the indexes were translated in English
 - 4.9% non-translatable titles: Austria (12.2%), Netherlands (7.4%), Sweden (9.9%)
more titles in national coding index, higher percentage of non-translatable entries ($r=.80$)
 - remaining: 60,559 records, of which 68% had no duplicate title
 - remaining: 19,190 records with in total 5,754 occupational titles (3.6 records per title)
 - of the 5,754 titles >> 3,131 have a 100% similar code across the indexes,
applying to slightly more than half of the titles (54%)



Database of translated job titles

- **In 2005 WageIndicator expanded its survey across countries**
 - The job title look-up table was translated
 - Gradually, more languages were added, and number of titles increased
- **>> WISCO database**
 - List of source titles was checked against coding indexes and by experts
 - Source titles could mostly be translated from English in national languages
 - If two source titles were translated similarly, the duplicate at highest skill level was removed, e.g. accountant vs bookkeeper
 - No translation if source title was not present in country, e.g. regional police officer
 - Translations were checked

Can surveyholders use the database?

- **Yes, free downloadable from website surveycodings.org**
 - Read the explanatory text, the papers and download the excel files
 - Use the live search
 - Use the API in your web-survey
 - Get in contact with Centerdata to discuss possibilities for CAPI software
- **What the database cannot do**
 - No mapping table to national occupational classifications (e.g. France, Germany, UK, Poland)
 - The excel file can be used for office coding, but so far no scripts available for removing typing errors and unidentified titles need to be coded manually
 - IER U Warwick offers its' CASCOT tool including several languages



Web surveys

- **Web surveys on laptops with internet connection**
 - Occupation API can be used, see surveycodings.org
 - The dataset will include an ISCO-code
- **Web surveys on smart phones**
 - preferably no search tree, only text box with short match list
- **Web surveys on an app without internet connection**
 - including the entire WISCO database in app is too much MBs
 - install selected parts of WISCO, e.g. country lists,
 - WageIndicator conducts surveys on an app, with reduced part of database, only including the occupational titles for targeted population

Database extensions - plans

- **Occupations not listed in database**
 - rare or new occupational titles
 - > 'suggest new item' box (to be developed) with office coding
- **Occupation question for other respondents than job holders**
 - what is occupation father/mother? -> response is at higher aggregation level
 - what occupation are you studying/looking for (students, job seekers)
 - -> WISCO database needs adaptation
- **Occupation – industry prediction**
 - measuring industry (NACE class) is as difficult as measuring occupation
 - we developed a prediction of the most likely industries based on occupation
 - respondents select their industry from a customized, reduced list, with 'other'
 - still to be programmed in the tool

Thank you for listening

- **Questions?** k.g.tijdens@uva.nl
- **Further reading**
- Tijdens KG (2014) Drop-out rates during completion of an occupation search tree in web-surveys, *Journal of Official Statistics*, doi.org/10.2478/jos-2014-0002
- Tijdens KG, De Ruijter E, De Ruijter J (2014) Comparing work tasks of 160 occupations across eight European countries, *Employee Relations*, doi.org/10.1108/ER-05-2013-0046
- Tijdens KG (2015) Self-identification of occupation in web surveys: requirements for search trees and look-up tables, *Survey Methods: Insights from the Field*, DOI:10.13094/SMIF-2015-00008
- Belloni M, Brugiavini A, Meschi E, Tijdens K (2016) Measurement error in occupational coding: an analysis on SHARE data, *Journal of Official Statistics*, doi.org/10.1515/JOS-2016-0049
- Tijdens KG, Kaandorp CS (2018) Validating occupational coding indexes for use in multi-country surveys, *Survey Insights: Methods from the Field*, DOI:10.13094/SMIF-2018-00007
- Tijdens KG, Beblavý M, Thum-Thysen A (2018) Skill mismatch comparing educational requirements vs attainments by occupation, *International Journal of Manpower*, doi/full/10.1108/IJM-10-2018-0328
- Tijdens, K.G. (2020). [Managing surveys: ten lessons learned from web-surveys](#)