



European Policy Brief no 3 November 2009

EurOccupations: CASCOT software for coding job titles

EurOccupations: Developing a detailed 7-country occupations database for comparative socio-economic research in the European Union

Finished FP6 project: April 2006 – April 2009

Content

EurOccupations: CASCOT software for coding job titles	1
Project summary	1
CASCOT software for coding job titles.....	2
Project information EurOccupations.....	3

Project summary

Occupation is the key unit in matching vacancies and job seekers, and it is used for occupational choice and for career consultancy. Occupation is a key variable in social research, particularly that which relates to the labour market, transitions from school to work, social stratification, gender wage gaps, occupational structures and skill requirements. The ILO's International Standard Classification of Occupations (ISCO) is the classification most commonly used, and it was recently updated to ISCO-08. Eurostat uses ISCO for its employment statistics. The EurOccupations project aimed to build a publicly available database containing the most frequent occupations for use in multi-country data-collection,

through the Internet or otherwise. It covered eight EU countries, notably Belgium, France, Germany, Italy, Netherlands, Poland, Spain, and United Kingdom. The database includes a source list of 1,594 distinct occupational titles within the ISCO-08 classification, country-specific translations and a search tree to navigate through the database.

CASCOT software for coding job titles

Data collectors in the UK and other countries use software programs for automated coding of texts. One example of such software is CASCOT (Computer-assisted Structured Coding Tool), which is developed and distributed by one of the EurOccupations partners, the Institute for Employment Research at the University of Warwick. So far CASCOT has mainly been used for coding job titles in surveys, for example classifying the reported occupations of the UK Labour Force Survey into the National Occupational Classification (SOC2000).

As part of the EurOccupations project, a multi-language coding tool has become available (Elias, Ellison and Jones 2009). CASCOT has been extended to facilitate the coding of occupational texts into the ISCO'08 classification. The software matches reported job titles in any of the seven languages of the participating countries to the EurOccupations database, resulting in a suggested ISCO'08 code for each job title. If, for example, the text *Hairdresser* is entered to the software, the best matching ISCO'08 code is shown. A certainty score (0-100) approximates the probability that the recommended code is correct. CASCOT allows the matching of Dutch, English, French, German, Italian, Polish and Spanish occupational texts, e.g. 'I am helping people with computer problems and questions', 'Ik ben administratief medewerkster' or 'Hôtesse d'accueil'.

Additionally the software acts as a 'look-up' tool: the EurOccupations database can be visited via the software by entering an occupational title in national language. The software provides a link to a related document which can be opened in a web browser. For all 150 key occupations, there is a detailed document available which contains information about occupational tasks, competencies, educational levels, etc.

CASCOT allows for fully manual, semi-automatic and fully automatic coding. Under normal circumstances, an experienced person can code manually about 100 occupations in an hour. If thousands of occupational texts are to be coded, the automatic mode of the software can be utilised. The texts are fed into the software in an input file. CASCOT will go through the file assigning ISCO'08 codes to all rows and saving the result to an output file for the user. Further information about CASCOT

<http://www2.warwick.ac.uk/fac/soc/ier/publications/software/cascot>

Project information EurOccupations

Funding scheme

FP6, Research Area 3.2.2:

Promotion and support for comparative research, methodologies and data generation
STREP PRIORITY 7, N° FP6-028987, Budget € 1,150,000

European Commission, DG Research, Project Officer Tihamer Laczko

Websites

www.eurooccupations.org

<http://eurocc.icares.com>

www.wageindicator.org

Further reading

Elias, P., R. Ellison, and R. Jones. 2009. "Development of associated software for EurOccupations database." EurOccupations Deliverable 26

For more information

k.g.tijdens@uva.nl or Peter.Elias@warwick.ac.uk

Coordinator

Prof.dr. Kea Tijdens, Erasmus University Rotterdam, Department of Social Sciences

Consortium

Netherlands [Erasmus University Rotterdam, Department of Social Sciences](#)

Netherlands [Arbeid Opleidingen Consult Tilburg](#)

United Kingdom [Warwick Institute for Employment Research, University of Warwick](#)

Belgium [Catholic University of Leuven, Hoger Instituut voor de Arbeid](#)

Poland [University of Economics Poznan, Katedra Pracy i Polityki Społecznej \(Unit of Work and Social Policy\)](#)

Netherlands [University of Amsterdam, Amsterdam Institute for Labour Studies \(AIAS\)](#)

Netherlands [Wage Indicator Foundation](#)

Spain [Fundació Centre d'Iniciatives i Recerques Europees a la Mediterrània \(CIEM\)](#)

France [Centre d'études et de recherches sur les qualifications \(CEREQ\), Marseille](#)

Germany [AWWW GmbH Göttingen](#)