



**SURVEY
FUTURES**
SURVEY DATA COLLECTION
METHODS COLLABORATION

Occupation Coding - Discussion

Survey Futures Workshop: Industry and Occupation Coding | University College London | 4 June 2026



Workshop Summary

- Look-up tools achieve high completion and face validity, but agreement with office coding and stability over time are only moderate
- Closed-list approaches reduce burden and cost – but consistency with office-based coding relatively low. The approach may be suitable for aggregate statistics but can lose nuance and confuse respondents
- Respondents often struggle to map jobs to formal SOC/SIC concepts; mental models rarely align with classification language
- Mode switching affects coding consistency
- LLMs show promise for both automated post-survey coding and real-time coding during data collection

Discussion Summary

- **Traditional coding remains effective**, but participants questioned whether it should continue to be viewed as the sole "gold standard" given variation across organisations, coders, and collection modes.
- **Look-up tools show potential but require further development** to improve usability and better reflect the language respondents use to describe their jobs and industries.
- **Closed-list approaches reduce burden and costs**, but respondents often struggle to identify with higher-level classification categories, potentially reducing coding accuracy and detail.
- **Respondent experience is critical**: many respondents find it difficult to map their occupation or industry to formal classification systems.
- There was broad agreement on the importance of developing **more respondent-friendly coding approaches** that align with how people naturally describe their work.

Discussion Summary

- **AI and Large Language Models (LLMs) were viewed as highly promising for both:**
 - Coding open-text responses after data collection.
 - Supporting coding during data collection through intelligent probing and follow-up questions.
- **AI has the potential to improve efficiency, consistency, and coding quality, but further validation and testing are needed** before widespread adoption.
- **Key considerations include:**
 - Transparency and explainability.
 - Data privacy and governance.
 - Model stability and long-term sustainability.
 - Industry-wide standards and best practices.
- **Continued research, collaboration, and evaluation across organisations will be essential to develop and assess these approaches.**



**SURVEY
FUTURES**
SURVEY DATA COLLECTION
METHODS COLLABORATION

Thank you for your attention!

