

Induced Automation: Classification of Automation Technologies

Current version: August 2022

This folder contains the automation innovation classification of 'Induced Automation: Evidence from Firm-level Patent Data' authored by Antoine Dechezleprêtre, David Hémous, Morten Olsen, and Carlo Zanella.

The analysis is performed using the 2018 EP full-text database of the European Patent Office on 1,538,370 patent applications with a description in English from 1978-2018. We compute the frequency of automation-related keywords for CPC and IPC codes at the main group level (referred to as the 6-digit level), for pairs of subclasses (referred to as 4-digit pairs), and for pairs of subclasses (i.e., 4-digit codes) with the classes G05 and G06. At the 6-digit level, we combine IPC and CPC codes, restrict attention to codes with at least 100 patents, and group 6-digit codes in catch-all categories ('xx' codes) at the 4-digit level when there are not enough patents. We call the resulting codes and pairs of codes 'technology categories'.

The folder contains the following files:

- `ipc6_cipc6xx_crosswalk.dta` gives a crosswalk from 6-digit IPC codes to our `cipc6xx` codes.
- `cpc6_cipc6xx_crosswalk.dta` gives a crosswalk from 6-digit CPC codes to our `cipc6xx` codes.¹
- `cipc6xx_classification.dta` gives the classification results at the 6-digit level.
- `cipc4_pairs_classification.dta` gives the classification results for pairs of 4-digit codes.
- `cipc4_G05G06_classification.dta` gives the classification result for 4-digit codes in combination with G05 or G06.

The variables are defined as follows:

- `total` total number of patent applications
- `cadcam` count of patent applications with at least one computer-aided design and manufacturing keywords
- `cnc` count of patent applications with at least one CNC keyword
- `plc` count of patent applications with at least one PLC keyword
- `automat` count of patent applications with at least one `automat*` keyword
- `flexman` count of patent applications with at least one flexible manufacturing keyword

¹uses `cpc-ipc-concordance.txt`, downloaded from the CPC website.

- `labor` count of patent applications with at least one labor keyword
- `robot` count of patent applications with at least one robot keyword
- `threedee` count of patent applications with at least one 3D printing keyword
- `anyclassification` count of patent application with at least one of the automation keywords (CNC, PLC, automat*, flexible manufacturing, labor, robot and 3D printing)
- `share_class` ratio of *class* patents over total patents, where *class* is the keyword classification categories above (cadcam, cnc, ..., anyclassification)
- `techn_field_nr` code for the technological field
- `techn_field` technological field
- `techn_sector` technological sector
- `machinery` dummy variable indicating whether this technology category is a machinery category (see definition in paper).
- `auto95` dummy variable indicating that the category is an auto95 automation technology
- `auto90` dummy variable indicating whether this category is an auto90 automation technology

For further details on the classification, please consult Appendix A.2 of the paper. For questions, feedback or inquiries, please write david.hemous@econ.uzh.ch. The classification is free to use and modify given attribution to the source (ODC-BY). The recommended citation is:

Dechezleprêtre, A., Hémous, D., Olsen, M. and Zanella, C. (2022). Induced Automation: Evidence from Firm-level Patent Data.