

NIOSH Industry and Occupation Computerized Coding System

May 2016

What are our priorities?

The National Institute for Occupational Safety and Health (NIOSH) NIOSH Industry and Occupation Computerized Coding System (NIOCCS) is a web-based system that translates industry and occupation text into standardized industry and occupation (I&O) codes. NIOCCS facilitates the coding of occupational data in order to increase the research and surveillance of occupationally related diseases. The system focuses on these areas:

- Increasing the percentage of data automatically coded.
- Maximizing the use of existing datasets for occupational safety and health research and surveillance.

What do we do?

- Design and develop new capabilities in the NIOCCS system to autocode occupation information in files where the industry code already exists, such as worker's compensation data.
- Conduct analysis to identify coding issues that can be solved by continual improvement to the system's artificial intelligence.
- Develop, test, and implement advanced pattern matching algorithms to improve NIOCCS autocoding accuracy.
- Improve awareness of NIOCCS through conference presentations, webinars, handout materials, NIOSH Science Blog, Wikipedia, publications, etc.
- Ensure that training, user support, and I&O coding consultation are accessible and available.

What have we accomplished?

- Increased the autocoding rate over time through continual improvements to the coding engine and knowledgebase.
- Processed over 14 million records as of January 2016 to NIOCCS for coding.
- Coded approximately \$2.5 million's worth of NIOSH surveillance and research project data as of January 2016.
- Successfully competed for NIOSH funds to modify NIOCCS to enable autocoding of occupation information in data files where industry codes exist, such as worker's compensation data.
- Provided I&O coding training to external users in Nevada, Georgia and New York in 2015; training planned for Kansas in June 2016.
- Registered over 400 people for a NIOCCS account as of January 2016. This user community comprises 225 unique organizations, including 36 state governments.

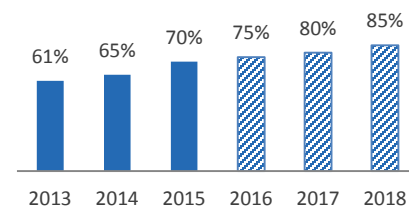
What's next?

- Increase autocoding performance and add new features to code occupation information in data files with existing industry codes.
- Update all system documents and training materials for the new version of NIOCCS.
- Implement the new version of NIOCCS for public use.
- Present the purpose, features, benefits, and impact of the NIOCCS system at one or more health-related conferences, such as, the National Association for Public Health Statistics and Information Systems (NAPHSIS), National Cancer Registrars Association (NCRA), Council of State and Territorial Epidemiologists (CSTE), Worker's Compensation Conference, etc.
- Provide information once implemented about the new version of NIOCCS on the NIOSH Science Blog.
- Develop targeted training videos that can be posted on the web.

At-A-Glance

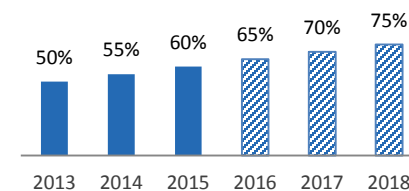
The NIOSH Industry and Occupation Computerized Coding System (NIOCCS) is a web-based system that translates industry and occupation text into standardized industry and occupation (I&O) codes to reduce the high cost and variability of manual coding. This snapshot shows recent accomplishments and upcoming work.

Percent of death certificate records autocoded with $\leq 10\%$ error rate



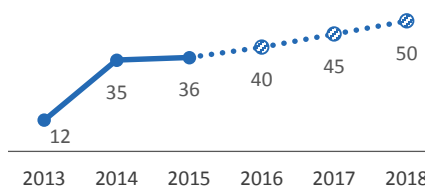
Source: NIOSH program records

Percent of all record types autocoded with $\leq 10\%$ error rate



Source: NIOSH program records

Number of states using NIOCCS to code industry and occupation



Source: NIOSH program records