

Industry and Occupation

CSTE DSWG Data Element Specific Team
Background Information



CSTE's Data Standardization Workgroup is working to define core data elements for transmission to CDC as a part of case-based surveillance for situational awareness or routine, cross-condition case surveillance. Industry and Occupation is a data class that contains elements to help describe the present or past work history of an individual.



This presentation will take a closer look at how various standards or agencies have defined data elements related to Industry and Occupation.

Potential data elements related to Industry and Occupation



- **Current Industry**
- **Current Occupation**
- Usual Industry
- Usual Occupation
- Occupation and Industry Category
- Occupational Duties
- Employment Status
- Work Site of Exposure
- Start Date
- End Date
- Food Handler
- Case Patient a Healthcare Worker
- Type of Healthcare Occupation
- Day Care Worker Within 7 Days Before Illness
- Visit Animal Environment Within 14 Days Before Symptom Onset
- Workplace Setting
- Critical Infrastructure Workplace Setting
- Employer address and name

*Bolded data elements included in MMG standardized Industry and Occupation Repeating Group

Council of State and Territorial Epidemiologists

Here is a brief overview of many of the data elements related to industry and occupation used either in CDC NNDSS message mapping guides (MMGs) or established in other standards. Current industry and current occupation, in bold, are likely to be the most familiar given they are present in the standard template used across many MMGs.

- Industry refers to the setting where someone works, such as restaurant, school, or hospital. Occupation refers to the kind of job someone has, such as pilot, nurse, or teacher.
- Data elements such as "food handler" and "case patient a healthcare worker" are used across multiple MMGs. There are also data elements that exist in other standards but are not used in any MMGs. An example is "Usual industry or occupation," which represents the type of work (paid or unpaid) done by a person for the longest amount of time during his or her life.

U.S. Core Data for Interoperability (USCDI)



Occupation Industry (Version 3)

Type of business that compensates for work or assigns work to an unpaid worker or volunteer. Examples include but are not limited to U.S. Army, cement manufacturing, and children and youth services.

- Value set: Industry NAICS Detail (ODH) & Industry CDC Census 2010

Occupation (Version 3)

Type of work of a person. Examples include but are not limited to infantry, business analyst, and social worker.

- Value set: Occupation ONETSOC Detail (ODH) & Occupation CDC Census 2010

Council of State and Territorial Epidemiologists

USCDI, or the US Core Data for Interoperability, is a standard published by the Office of the National Coordinator for Health Information Technology that aims to standardize a set of health data classes and constituent data elements for use in nationwide, interoperable health information exchange.

Starting with version 3 of USCDI, both occupation industry and occupation were included as data elements within the Patient Demographics/Information class. What is referred to as “Occupation Industry” in USCDI and in the HL7 templates we will speak about in a few minutes is equivalent to “Industry” in the MMGs.

Two value sets are listed as applicable standards in USCDI for each of these elements. NAICS and CDC Census 2010 for Industry and ONETSOC and CDC Census 2010 for Occupation.

- NAICS, the North American Industry Classification Scheme, is a hierarchical structure for classifying industry that was developed by the US government’s Office of Management and Budget. It is used by federal agencies to classify business establishments and is updated every five years.
- SOC, the Standard Occupational Classification, is developed and maintained by the U.S. Bureau of Labor Statistics. It is used by federal agencies to classify workers

into occupational categories and is updated every five years. O*NET (the Occupational Information Network) has an extension of the SOC system from over 800 classifications to more than 1100 classifications. This O*NETSOC is a hierarchical structure as well.

It is important to note that there is no relationship between the codes in NAICS and SOC.

Census industry and occupation codes are derived from NAICS and SOC and are maintained by Census Bureau. These Census industry and occupation codes can be related or paired. The Census codes are less detailed than NAICS and SOC to protect privacy.

U.S. Core Data for Interoperability (USCDI)



Usual Industry (Level 2)

A self-reported term that identifies the kind of business, i.e., primary business activity, in which a person has worked for the longest time while in their Usual Occupation. For a military position, this is the self-reported branch of service. If an appropriate term is not available (e.g., a new kind of business), then a text entry is used.

- Value set: Industry NAICS Detail (ODH) & Industry CDC Census 2010

Usual Occupation (Level 2)

A self-reported, coded term for the type of work (paid or unpaid) done by a person for the longest amount of time during his or her life, not including voluntary work (done by choice for the benefit of others without compensation)

- Value set: Occupation ONETSOC Detail (ODH) & Occupation CDC Census 2010

Employment Status (Level 1)

This data element includes a coded concept, Employment Status, with time and date stamp and/or start date and end date. A person can be both employed and retired, and can have a different employment status for each job. Employment Status is a person's self-reported, coded relationship to working for pay, family earnings, or training (e.g. having one or more jobs, searching for work, etc.). A person's Employment Status is independent of Job characteristics, e.g., not "full-time work" or "part-time work," because many people have more than one job.

Council of State and Territorial Epidemiologists

In addition to occupation and occupation industry, three additional data elements exist at either level 2 or level 1 for consideration for adoption within future versions of USCDI. The data elements classified as Level 2 have broad standardization and guidance for their collection, and there are existing terminology standards or technical specifications for their representation. They are already being captured and exchanged in multiple EHRs and can be applied to most care settings. By comparison, Level 1 data elements are less mature but are still represented by an existing terminology standard or technical specification, are captured and exchanged in at least one EHR, and can be applied to several care settings.

"Usual industry" and "usual occupation" are level 2. These data elements capture the industry and occupation where someone worked for the longest amount of time during his or her life. These have the same value sets as occupation industry and occupation seen on last slide.

"Employment status" is at level 1. This is a person's self-reported, coded relationship to working for pay, family earnings, or training.

HL7® v2

- I/O was introduced in v2.9 using new segments OH1-OH4.
 - OH1 (person employment status)
 - Employment status, start date, end date
 - OH2 (past or present job)
 - Example:

Occupation

OH2[1|A|20230707|4030^Food preparation workers^CDCOCCUPATION2010|4970^Grocery stores^CDCINDUSTRY2010|PWSE^Paid work, self-employed^HL70959|20200707|||10|7||||Packing food for delivery|||Y^Yes^HL70136

Job Duties

Industry

Current job Indicator

Work classification
 - OH3 (usual work)
 - Occupation, industry, duration, start year
 - OH4 (combat zone work)
 - Start date and end date

Council of State and Territorial Epidemiologists

NIOSH has been working to establish standards and tools to support the collection and use of self-reported Occupational Data for Health (sometime abbreviated “ODH”). They have been working to have this information included in key data-formatting interoperability standards for health information exchange, notably HL7 v2, CDA, and FHIR.

In HL7 v2.9, industry and occupation were introduced using the segments OH1-OH4:

- OH1 captures the employment status of the patient, and the start and end date of that employment status. This segment can repeat to indicate different employment status designations over time for a patient.
- OH2 captures either a past or present job of the patient. This segment can also repeat to enable sending the job history of a patient over time. This is particularly relevant for case notification data, as this segment is used to communicate the information about the job or jobs which the subject currently holds or has held in the past. It includes industry and occupation, start and end dates of the job, and a current job indicator.
- OH3 captures the usual occupation and industry, which is what the patient has held for the longest duration throughout their life. In this segment industry and occupation are present, as well as the duration of their usual occupation in years.

Only one usual occupation and industry is permitted for an individual.

- OH4 captures the date range for when a patient has worked in what was considered a combat or hazardous duty zone, either civilian or military.

The V2.9 ODH segments may be retrofitted into older versions of HL7 v2 for specific purposes, or included when the newer version is adopted.

HL7® Clinical Document Architecture (CDA)



- NIOSH Occupational Data for Health (ODH) Templates are in some CDA implementation guides
 - Electronic Initial Case Reporting (eICR)
 - Cancer Registry
 - Vital Records Death Reporting
- CDA Mappings
 - Past or Present Occupation Observation
 - **History of Occupation Industry (86188-0)**
 - **History of Occupation (11341-5)**
 - Other observations (Job duty, Occupational Hazard, work schedule, work classification, supervisory level)
 - Usual Occupation Observation
 - **History of Usual Industry (21844-6)**
 - **History of Usual Occupation (21843-8)**
 - Usual occupation duration (74163-7)
 - History of Employment Status Observation
 - **History of employment status NIOSH (74165-2)**
- Templates also support inclusion of Combat Zone Period and Date of Retirement Observations

*Bolded data elements used in eICR
Council of State and Territorial Epidemiologists

For HL7 Clinical Document Architecture, or CDA, NIOSH has supported the development of templates that can be used to include occupational health data in the social history section of any CDA document.

Of importance to us, the templates support collection of observations on present (and past) industry and occupation. Also present in the template are employment status, usual industry and occupation, and combat zone period, similar to what was discussed on the previous HL7 v2 slide.

So far, the Occupational Data for Health templates have been included in implementation guides for electronic initial case reporting, cancer registries, and vital records death reporting. The bolded fields in the slide represent those included in eICR Implementation guide.

HL7® Fast Healthcare Interoperability Resources (FHIR)



- Occupational Data for Health (ODH) FHIR IG
 - Uses the same content in the CDA ODH Templates.
- Resource Profiles
 - Past Or Present Job
 - History of Occupation Industry (86188-0)
 - History of Occupation (11341-5)
 - Other observations include Job duty, Occupational Hazard, work schedule, work classification, supervisory level)
 - Usual Work
 - History of Usual Industry (21844-6)
 - History of Usual Occupation (21843-8)
 - Usual occupation duration (74163-7)
 - Employment Status
 - History of Employment Status NIOSH (74165-2)
 - Other profiles include Combat Zone Period and Retirement Date

Council of State and Territorial Epidemiologists

For FHIR, the ODH Implementation Guide is a composition of profiles on the observation resource. The scope of work-related information aligns with CDA as it is similarly based off of the ODH information model. It includes resource profiles for past or present job, usual work, employment status, retirement date, and combat zone period.

Case Notifications to CDC Message Mapping Guides (MMGs)



- Template included in Carbon Monoxide, TB/LTBI, Rubella, Congenital Rubella, Measles, RIBD, Babesiosis, FDD, HAI MDRO, COVID, Hepatitis v2, Malaria, Mumps, Trichinellosis, and Varicella MMGs

Current Occupation (85658-3)	This data element is used to capture the narrative text of a subject's current occupation.	Text
Current Occupation Standardized (85659-1)	This data element is used to capture the CDC NIOSH standard occupation code based upon the narrative text of a subject's current occupation.	Occupation CDC Census 2010
Current Industry (85078-4)	This data element is used to capture the narrative text of a subject's current industry.	Text
Current Industry Standardized (85657-5)	This data element is used to capture the CDC NIOSH standard industry code based upon the narrative text of a subject's current industry.	Industry CDC Census 2010

Council of State and Territorial Epidemiologists

The template of four data elements seen in the table on this slide is currently used across many message mapping guides to collect information on occupation and industry. The same occupation may be found across multiple industries, so it is valuable to have both pieces of information to get a better understanding of occupational risk and exposures. For example, a nurse in a hospital and in a school are likely to have different potential exposures.

The list on the top of the slide indicates which MMGs this template is currently part of. Jurisdictions can send both narrative text (in a text data element) or a standardized census code (in a coded data element). This template in the NNDSS MMGs currently uses the value sets from the 2010 census codes for industry and occupation.

It is recommended that jurisdictions accept and send the free text and coded values when they are available, but should still accept and send free text if the codes are not available. Jurisdictions can run the free text thru NIOCCS for coding before sending it to CDC, but if this isn't possible, jurisdictions can send the free text values for CDC to code through NIOCCS after receipt.

- NIOCCS stands for the NIOSH Industry and Occupation Computerized Coding

System. Briefly, NIOCCS is a free web-based API that converts industry and occupation text into standardized codes and can output codes from the Census coding schemes, or NAICS and SOC. More information on NIOCCS is available in the April 2023 eSHARE call recording, available in the NNDSS Technical Resources Center on CDC's website.

Case Notifications to CDC Message Mapping Guides (MMGs)



- Rubella, Measles, Pertussis, Varicella, COVID-19, Hepatitis v2

Data Element Name	Data Element Description	Value Set
Case Patient a Healthcare Worker	Hep A: Was the patient employed as a healthcare worker during the TWO WEEKS prior to onset of symptoms to ONE WEEK after onset of JAUNDICE? (If no jaundice, use two weeks after onset of symptoms.) Others: Was patient healthcare personnel (HCP) at illness onset?	1. Yes 2. No 3. Unknown

Council of State and Territorial Epidemiologists

Several NNDSS message mapping guides ask specific questions about the case patient working in a particular occupation or industry. The most common asks whether the case patient was a healthcare worker. This is used in 6 guides.

Prior work on Industry and Occupation



NIOSH Guidance

- Use the templates provided in HL7 v2.9, CDA, and FHIR resources
- Collecting
 - Both, Industry and occupation, should be collected as they are related, but different
 - Collect free text descriptions; do not use drop down lists
 - Collect detailed information
 - Each occupation and industry has a specific standardized code assigned to analyze the data
- Coding
 - Industry and occupation codes are standardized
 - Review/decide the classification systems to use before coding data
 - Use an autocoder to code your data, such as NIOCCS
 - You can code data as you collect it, or after collection is complete

Council of State and Territorial Epidemiologists

In general, NIOSH recommends use of the HL7 templates provided for v2.9, CDA, and FHIR data exchange, discussed previously.

Important things to remember when collecting industry and occupation data include the following:

- Industry and occupation are related, but different. It is important to collect both.
- It is important to collect free text descriptions and not use drop down lists – limiting the response options also limits what we can learn. Drop downs make it difficult to capture emerging occupations/industries.
- It is best to collect detailed information. For instance, “Engineer” is okay, “Civil Engineer” is better.
- Each occupation and industry has a specific standardized code assigned to it, and using these codes assists in data analysis, integration, and interpretation.

There are also important things to remember when coding industry and occupation data:

- It is important to decide which classification systems to use before coding the data. Descriptions of the code systems can be found on NIOSH's website.
- Critically, use an autocoder, such as NIOCCS, to code the data.

- Software applications, or “autocoders,” assign industry and occupation codes to free-text descriptions using the classification system you choose (NAICS and SOC, or Census).
- There are a number of benefits to using an autocoder:
 - Autocoding is much faster than manual coding.
 - There is greater consistency in codes from one record to another.
 - Autocoding does not require the extensive training and experience needed for accuracy with manual coding.
- Finally, data can be coded as it is collected, or as a batch after collection is complete.

Prior work on Industry and Occupation



Advantages of NAICS/SOC over Census codes

- Broader use
 - NAICS/SOC are more broadly used than the Census codes
- Hierarchical and detailed structure
 - Hierarchical nature of NAICS/SOC provides additional information
 - NAICS/SOC codes offer more detailed classification
- Clarity and independence in classification
 - Independence of NAICS/SOC providing clear and unambiguous coding
- Simplified mapping and crosswalk
 - Mapping from NAICS/SOC to Census codes is simple

Council of State and Territorial Epidemiologists

NIOSH also recommends the use of NAICS and SOC over use of the Census codes for a number of reasons:

- NAICS/SOC are more broadly used than the Census codes as standards for coding industry and occupation data.
- They are hierarchical while the Census codes are not, and this hierarchy adds useful information about relationships.
- NAICS/SOC provide more detailed information than the Census codes, which are intentionally less detailed to protect privacy.
- NAICS/SOC are unrelated, independent classification systems – so coding is clearer in cases when only industry or only occupation is available.
- And finally, mapping (or cross-walking) NAICS/SOC to Census codes is a simple process. NAICS/SOC are the foundation of the Census codes. Mapping from the more detailed NAICS/SOC to the less detailed Census codes can be done when it is helpful.

Prior work on Industry and Occupation



- Wisconsin DHS example
 - 27 conditions in WEDSS have I/O in their form
 - Use NIOCCS to translate text to codes
 - By January 2023, 87% of COVID-19 cases with completed interviews had complete industry and occupation data
 - Generally, I/O data is the only reliable way to identify, intervene, and prevent immediate and long-term workplace health hazards

Council of State and Territorial Epidemiologists

In Wisconsin, 27 conditions have industry and occupation in their surveillance system. The most recent conditions to have this added include: Anaplasmosis, Legionellosis, and Spotted Fever Group Rickettsiosis

They use NIOCCS to translate industry and occupation text fields into standardized codes.

- In January 2022, 47% of COVID 19 cases with completed interviews had complete industry data and 66.5% had complete occupation data.
- By January 2023, 87% of COVID 19 cases with completed interviews had complete industry and occupation data.

Wisconsin has found that industry and occupation information is essential for identifying, responding to, and preventing workplace health hazards.

For discussion



- Is the standard template of a repeating group for I/O present in 15 MMGs working?
- Are jurisdictions using NIOCCS to code text data?
- Should the value sets be updated from the 2010 CDC Census to NAICS/SOC?
 - 2017 NAICS (North American Industry Classification System) for industry codes
 - 2018 SOC (Standard Occupational Classification) for occupation codes
 - If code systems are changed, update the data element identifiers with the LOINC values for NAICS (for industry) and SOC (for occupation).
- Can we recommend a template to capture particular occupations across MMGs? (E.g., Patient is a healthcare worker, food handler, etc.)

Council of State and Territorial Epidemiologists

Here are some topics to think about for discussion before our group meeting on October 20th.

- Thinking about the standard template used across MMGs for Industry and Occupation- how is that working? Are there issues with collecting or sending current industry and occupation?
- Are jurisdictions using NIOCCS to code free text data?
- Should the value sets be updated from Census codes to NAICS and SOC code systems?
- Can or should the group recommend a template to capture particular occupations across MMGs in the same format? (For instance, If the patient is a healthcare worker or food handler.)