Injury Tracking Application (ITA) Case Detail Data Dictionary

This Case Detail Data Dictionary provides information about the fields in the publicly available ITA Case Detail dataset. Each record in the Case Detail dataset has information about a single work-related injury or illness. Standard Occupation Codes (SOC) have been added by OSHA to each record using the NIOSH Industry and Occupation Computerized Coding System (NIOCCS). At the end of each calendar year, Occupational Injury and Illness Classification System (OIICS) codes are added to each record when there is sufficient detail in the record and then made available on OSHA's website, typically by the end of January (e.g., the calendar year 2025 data is collected in 2026, and the final complete year of data with OIICS codes added will be posted by January 2027).

Data Element	Definition	Format	Source
ID	A unique number for each record.	Numeric	ITA assigned
establishment_ID	A unique number for each establishment. This field can be used to link establishment 300A data to 300/301 data.	Numeric	ITA assigned
establishment_name	The name of the establishment reporting data (must be unique to the establishment).	Character	OSHA Forms 300A, 300, 301
ein	Employer Identification Number (EIN), also known as Federal Tax Identification Number.	Character	OSHA Form 300A
company_name	The name of the company that owns the establishment.	Character	OSHA Form 300A
street_address	The street address of the establishment.	Character	OSHA Form 300A
city	The city where the establishment is located.	Character	OSHA Form 300A
state	The state or territory where the establishment is located.	Character	OSHA Form 300A
zip_code	The zip code of the establishment.	Character	OSHA Form 300A
naics_code	The North American Industry Classification System (NAICS) code for the establishment (can be 2012, 2017, or 2022).	Numeric	OSHA Form 300A
naics_year	The calendar year reflecting the version of NAICS codes used by the establishment. When a NAICS code has not changed, the most recent NAICS year is listed.	Numeric	ITA assigned
industry_description	The industry description for the establishment.	Character	OSHA Form 300A

Data Element	Definition	Format	Source
establishment_type	• 1 = Private industry • 2 = State government entity • 3 = Local government entity	Numeric	OSHA Form 300A
Size*	The size of the establishment is based on the employer-reported **maximum number** of employees who worked there **at any point** in the year. • 1 = less than 20 employees • 2 = 20-249 employees • 21 = 20-99 employees • 22 = 100-249 employees • 3 = 250+ employees	Numeric	OSHA Form 300A
annual_average_employees	The annual average number of employees at the establishment. Note: This field should not be summed across cases in an establishment.	Numeric	OSHA Form 300A
total_hours_worked	The total hours worked by all employees at the establishment. Note: This field should not be summed across cases in an establishment.	Numeric	OSHA Form 300A
case_number	An employer-assigned case number for each unique case (i.e., injured/ill employee).	Character	OSHA Forms 300 & 301
date_of_incident	The date the incident occurred.	Numeric	OSHA Forms 300 & 301
incident_outcome	The most serious outcome that occurred. • 1 = Death • 2 = Days away from work (DAFW) • 3 = Job transfer or restriction (DJTR) • 4 = Other recordable case	Numeric	OSHA Form 300
dafw_num_away	The number of days away from work the employee required to recover from the incident.	Numeric	OSHA Form 300
djtr_num_tr	The number of days the employee needed to be transferred or reassigned to another job or placed on restricted duty due to the incident.	Numeric	OSHA Form 300
type_of_incident	The type of incident that occurred. • 1 = Injury • 2 = Skin disorder • 3 = Respiratory condition • 4 = Poisoning • 5 = Hearing Loss • 6 = All other illness	Numeric	OSHA Form 300

Data Element	Definition	Format	Source
time_started_work	The time the affected employee started work prior to the incident, shown as HH:MM.	Numeric	OSHA Form 301
time_of_incident	The time the incident occurred, shown as HH:MM.	Numeric	OSHA Form 301
time_unknown	Was the time of the incident unknown? • 0 = No • 1 = Yes	Numeric	OSHA Form 301
date_of_death	The date the death occurred, if applicable.	Numeric	OSHA Form 301
created_timestamp	The date and time the record was submitted to the ITA.	Numeric	ITA assigned
year_of_filing	The calendar year in which the injuries and illnesses being reported occurred at the establishment.	Numeric	OSHA Form 300A
job_description	The job title of the employee.	Character	OSHA Form 300
SOC_code[1]	The 2018 SOC assigned by NIOCCS.	Numeric	NIOCCS assigned
SOC_description[1]	Text description of the 2018 SOC Code.	Character	NIOCCS assigned
SOC_probability[1]	The score given by NIOCCS for the accuracy of the SOC code. The score ranges from 0 to 1.0; the higher the score, the higher probability the SOC has been accurately identified.	Numeric	NIOCCS assigned
Unexpected_NAICS_SOC_Combo[1]	Indicates industry/occupation code combinations that are unlikely to be used together according to data collected by the U.S. Bureau of Labor Statistics. • Y = Yes • N = No	Character	NIOCCS assigned
New_incident_location[2]	Where the event occurred.	Character	OSHA Form 300

Data Element	Definition	Format	Source
New_incident_description[2]	Description of the injury or illness, parts of body affected, and object/substance that directly injured or made the person ill.	Character	OSHA Form 300
New_nar_before_incident[2]	What was the employee doing just before the incident occurred?	Character	OSHA Form 301
New_nar_what_happened[2]	What happened? Tell us how the injury occurred.	Character	OSHA Form 301
New_nar_injury_illness[2]	What was the injury or illness?	Character	OSHA Form 301
New_nar_object_substance[2]	What object or substance directly harmed the employee?	Character	OSHA Form 301
Nature_code_pred[3]	The code for the principal physical characteristic(s) of the work-related injury or illness.	Numeric	BLS assigned
Nature_title_pred[3]	A text description of the assigned nature of injury or illness code.	Character	BLS assigned
Part_code_pred[3]	The code for the part of body directly affected by the identified nature of injury or illness.	Numeric	BLS assigned
Part_title_pred[3]	A text description of the assigned part of body code.	Character	BLS assigned
Event_code_pred[3]	The code for the event or exposure that identifies the manner in which the injury or illness was produced or inflicted.	Numeric	BLS assigned
Event_title_pred[3]	A text description of the assigned event code.	Character	BLS assigned
Source_code_pred[3]	The code that identifies the objects, substances, equipment, or other factors that harmed the worker.	Numeric	BLS assigned
Source_title_pred[3]	A text description of the assigned source code.	Character	BLS assigned

Data Element	Definition	Format	Source
Sec_source_code_pred[3]	The code, if any secondary objects, substances, equipment, or other factors that harmed the worker.	Numeric	BLS assigned
Sec_source_title_pred[3]	A text description of the secondary source, if any.	Character	BLS assigned

*From 2016-2022, the size field had three options (1 = <20, 2 = 20-249, 3 = 250+). In 2023, to accommodate for the 300/301 submission requirements for establishments with 100 or more employees, the second category was split into 21 = 20-99 and 22 = 100-249. However, some establishments may still continue to show a size coding of "2".

[1] OSHA used NIOCCS to assist in assigning SOC codes using job_title and NAICS_code. The NIOCCS system is a free web-based application that uses machine learning, a form of Artificial Intelligence (AI), used to assign SOC codes, SOC descriptions, and SOC probability for ITA cases submitted. NIOCCS also generates a flag if the SOC code and industry code are unlikely combinations to support additional validation checks. Note that OSHA does not validate the NIOCCS output. If NIOCCS could not assign a SOC code (such as due to vague job description, low NIOCCS probability, or NIOCCS determining there was insufficient information to assign a code) these entries were assigned a SOC code "00-9999". For more information about the NIOCCS-generated codes see NIOSH's "About NIOCCS" page.

[2] OSHA used AI and some manual review to detect and remove information that could reasonably be expected to identify individuals directly from the narrative, open-text fields before making that data publicly available.. In most cases, only a portion of the text is redacted; in some cases an entire field has been redacted. The prefix "New_" was added to the original data element name to indicate the redaction of personally identifiable and sensitive medical information.

[3] OSHA collaborated with the Bureau of Labor Statistics (BLS) to use the Survey of Occupational Injury and Illness (SOII) autocoder to assign Occupational Injury and Illness Classification System, Version 2.01 (OIICS) codes for nature of injury, event or exposure, part of body, source of injury, and secondary source of injury based on information in the six ITA narrative fields. OIICS codes are organized as a four-digit tree-like hierarchical coding structure with parent-child relationships. However, OIICS codes assigned by BLS to the OSHA ITA data are limited to only two-digit OIICS codes. Furthermore, only non-terminal two-digit OIICS codes are assigned. OIICS non-terminal codes are those with three- or four-digit child codes. In the OIICS Version 2.01 Manual, these codes are marked with an asterisk. Some two-digit OIICS codes may also be terminal, meaning that they do not have three- or four-digit child codes; these terminal OIICS codes were not assigned and the data element will be blank. Out of 289 possible two-digit codes, 203 are non-terminal codes, which accounts for about 70% of the codes in the OIICS Version 2.01 Manual. While two-digit codes are less detailed than four-digit codes, they still allow for grouping of similar cases. Once grouped, data users can use the detailed narratives for analysis to identify trends and patterns in workplace injuries and illnesses, which can inform safety measures.