# Occupational License Review and Processing Summary

6.3.2025

**1. Data Intake**

States submit licensing data using two core tables:

* license: individual license records
* licauth: licensing authority information

Submissions are received in multiple formats (.xlsx, .csv, .accdb) through email or online upload, depending on each state’s infrastructure, historical practices, and regulatory constraints.

All submissions are downloaded to a centralized intake folder and clearly labeled by state, ensuring traceability and organization.

**2. Standardization and Coding**

Each state’s data is processed independently, replacing the previous vintage in full. This stage involves:

* **Integrity and Structure Checks**

Although states submit data in the correct license/licauth table structure, some fields can break when loading into the database because often the source is not configured in the same way (or is not a database at all). To get to the database structure and deliver the data in a reasonably consistent way to CareerOneStop, validation happens on a number of features.

* + Verifying correct state FIPS codes
  + Leading zeroes are evaluated (in CSV transfer formats these sometimes go missing and break id continuity)
  + Ensuring primary key and required fields contain data
  + Ensuring foreign key relationships between tables are valid and complete
  + Length of fields
  + Department/Division/Board titles sometimes get rearranged
  + Structured indicator fields get recoded if they’re using null instead of 9 or extra values
* **Occupational Coding**

To make occupational coding useful across states occupational codes need to be assigned consistently between states. Licenses can be very different from occupations and guidance from OEWS isn’t always the most helpful, so in the past there has been wide variation in state-assigned codes. For example, some licenses (doctors, teachers, engineers) consistently require multiple codes to be assigned, and some ambiguous licenses may be assigned to very different codes and states would use either SOC or ONET. For this reason and to reduce the burden on states, the requirement to submit licxocc content (assigned occupational codes) was removed several years ago and this is now done centrally.

The process for central coding was originally a custom autocoder that still exists and could be reverse-engineered (sproc\_AC\_insert in the License database of lmisqlstage\lmisqlstage), but the primary focus is now on continuity with previous vintages of licenses. That is done in steps.

* + Noting what identifier can be used to link current and prior vintage data
  + Retaining existing O\*NET codes where records remain unchanged
  + Coding new or modified licenses to the current O\*NET taxonomy
* **Integration of Additional Sources**

There are ways to systematically enrich the data to make it more complete in terms of depth or coverage.

* + **Centrally collected:** These are licenses that have been identified manually. Some occupations are nearly universally licensed and when they’re missing from a state the ARC has collected them separately and stored them in a parallel table to be merged later. (CC\_License, CC\_LicXOcc, CC\_LicAuth). These occasionally need to be disabled as states add them to their own data.
  + **Other data sources:** In 2017 we undertook a major overhaul to look for gaps in our coverage and we did that in part by looking at other data sources. Each of these privately developed sources had limitations in scope or coverage but was more complete for its area of focus, so combined could add either detail or licenses to our data set. Although state submissions have improved dramatically and most of these gaps are no longer being filled by the external data, the processes are still in place to add either new licenses or missing indicator field values from other sources. This is likely unnecessary.
  + **Inferred industry:** Some occupations are licensed only when working in specific industries. For example, states may require everyone who works in a casino to hold a registration. Bartenders in casinos then need to be licensed, but not elsewhere and that can throw off analysis of the share of workers who may need a license. Major categories like this have been identified (likely incomplete) and are noted in the licxind view.
  + **License compact status:** A number of private organizations have created compacts to help make specific licenses more transferrable. While many are still in process and not yet enforced, tracking these developments helps to see what state priorities are and some of the reciprocity efforts that are out there.
* **Composite Table Construction**
  + A unified national table is built, combining all validated sources and enriched content

**3. Export and Publication**

The finalized composite dataset is exported and published through the CareerOneStop (COS) platform. This ensures the data reaches a broad audience including jobseekers, career counselors, researchers, and policymakers. COS also integrates the data into O\*NET, expanding its impact across multiple workforce tools.

[wfinfodb/License|data.widcenter.org](https://data.widcenter.org/wfinfodb/License/)

Tables containing individual state data are also released for states that prefer to edit our data than start from scratch.

**4. National Analysis**

Many states review their licensing data in relation to specific policy concerns—such as military transferability, treatment of criminal records, fee structures, or reciprocity—but often lack the capacity or mandate to explore broader, cross-state trends. The ARC (Analyst Resource Center) is well-positioned to provide this kind of national perspective.

National analysis efforts serve to:

* **Support Policy and Research**: Identifying commonalities, outliers, and gaps across states provides insight into licensing practices and workforce implications.
* **Improve Data Quality**: Comparative analysis surfaces anomalies and informs targeted support to states.

For example, a dashboard comparing license coverage across states has been developed and maintained by the ARC. North Carolina has cited its utility in internal decision-making, and continued investment in such tools helps states benchmark their systems and identify improvement opportunities.

While longitudinal research using license data is still underutilized, it holds promise for understanding licensing trends over time and should be further explored.

**5. Promotion and Support**

States operate in dramatically different administrative and political environments, which affects their ability to collect and report license data. In some cases, the data is centrally managed and readily available. In others, it may reside with numerous fragmented and under-resourced licensing boards, be scraped from public websites, or gathered through direct outreach or survey-like processes.

When states seek to launch or improve their data collection efforts, they frequently turn to the ARC for guidance. Support can include technical advice, feedback on data structure, and examples of successful strategies from other states.

Beyond technical assistance, the licensing dataset requires **ongoing promotion and support** to remain relevant and widely used:

* **Engagement with the Public and Stakeholders**: CareerOneStop and O\*NET serve as key distribution channels, making the data accessible to a national audience.
* **Responsive Support**: The ARC answers questions from users, addresses issues, and facilitates new applications of the data.
* **Data Collection Support**: The ARC answers questions from states who are collecting the data and defining their processes.
* **DOL Priorities**: The ARC maintains an active relationship with the U.S. Department of Labor, responding to inquiries and showcasing the dataset’s value and alignment with workforce priorities.

This ongoing engagement helps keep the product fresh, visible, and aligned with both user needs and national goals.