Analyst Resource Center
Serving the workforce data community

Annual Report 2022

Summary of Accomplishments

LEWIS: Software development of a web-based version and MB3 method has been completed. One in-person training has occurred and more trainings are being planned.

WID 3.0: The structure is complete but we're awaiting feedback, first from ARC members then other state LMI offices.

Occupational Licenses: We're receiving regular submissions from the majority of states and are considering options to improve file transfer, including possibly an administrative tool.

Employer Database: DataAxle, the current provider of the data has offered us a number of enhancements which we've evaluated as needed.

Real-time Analytics: We work with the NLX Research Hub and CareerOneStop to evaluate access to NLX job postings data and help democratize its use.

Training: We're hosting a virtual Python Users Group and are hoping it continues to evolve and become established as a chance for states to expand their teams' base of technical skills.

Outreach: Two new states will be participating in the Spring 2023 ARC meeting, new members from existing states will be attending, and ARC staff have participated in a number of virtual and in-person national LMI events.

Documentation: One of the ways we serve states is by documenting potential WID content to help states find a data source that meets their needs. Traffic to our documentation makes up most of our website usage and two topics in particular drive that traffic – CIP-SOC crosswalks and Tableau data connectors. While in both cases we're mostly directing users to other sources, it gives us insight into topics that are of interest to the public and may not be well-documented elsewhere.

Contents

The Analyst Resource Center (ARC) is a cooperative group of states that comes out of the Workforce Investment Grant (WIG) funding activities ranging from implementation of the Workforce Information Database (WID), state and regional projections, occupational licenses, to research and publications work. The ARC's primary activities are in setting standards for the WID and in providing resources in support of grant activities that may be useful to a range of recipients. We hold twice-annual meetings with volunteer state participants to make decisions, and most of our resources are published on a website. 2020 was a tumultuous year. Many of our committee members were overwhelmed with new tasks and our traditional avenue of connection, in person meetings, was derailed. Despite this we were able to keep tabs on ongoing priorities and pivot to a digital format for many activities. This report is a summary of those activities and the ones we hope to keep up with going forward.

The ARC exists to help states achieve their data goals, centralizing some of the work of keeping systems up and going. Our efforts often center around the Workforce Information Database (WID), a database structure states implement that allows common definitions and use of data common in the LMI world. Having the common structure:

- Gives states that experience unexpected turnover to have a central resource that can help them make sense of priorities
- Allows improved communications between states
- Allows partner organizations that provide outputs to have a single point of contact for requirements rather than 50 different stakeholders
- Saves state resources from doing design work. Designing a database structure so that it can adapt to changes, include all appropriate details, and documenting that design so that others can use it is challenging. Many states don't have the staff to do that front-end work well and having the structure and documentation ready-made saves on problems that may crop up later

We also procure the Employer Database – a list of employers from a private vendor that allow users to find actual company names (something prohibited by CIPSEA rules for our own products) which would otherwise be available only to a few states, manage the LEWIS system (a piece of software that allows states to securely aggregate OES data for different regions or subsets of jobs than are covered by BLS publication). More recently we've worked on APIs and the challenges states face in either using those provided or in developing their own – again saving the difficult and error prone work of design.

While this type of work is a critical foundation for the management of LMI data in states, design work tends to happen at the front end of a project and the benefits are realized much later. The challenge is in staying ahead of state priorities so we already have structure and taxonomy to offer when states are looking for them and in ensuring that they know those resources are there even if they only need them every several years.

Committee actions

ARC Meetings

In 2022 ARC hosted its first in-person meeting since 2019. The meeting was held in Denver in late July. Most states that had historically participated returned, except for Connecticut, Wisconsin, and South

Carolina. South Carolina has participated intermittently due to turnover and Connecticut and Wisconsin were unable due to health or travel concerns.

The major topic of conversation was how to structure the organization going forward. During the preceding few years many of our major participants have retired or set a date to retire. LMI Directors and people in management roles in particular have departed, with the majority of attendees filling more technical roles within their organizations. Participants agreed to recruitment efforts.

Organizational Outreach

Subsequent to the ARC meeting in July, Bill McMahon and Amanda Rohrer have participated in a number of broader LMI events as presenters and participants. Many of those stemmed from the NLX Research Hub and a role in a user group. We presented about that at a conference in Chicago in the fall, and virtually we attended several other meetings throughout the year. During those meetings, the invitation to participate in ARC activities was issued and states expressed interest. As a result, for the upcoming meeting in spring 2023 we're expecting to have two new states represented and several new attendees.

Training

In December of 2022 a basic LEWIS training was held in Minnesota. Results from evaluations were all largely positive. There were 18 responses, 10 of which were all 5s with no actionable comments. The remaining 8 were also positive (all 4s and 5s, one person did 3s and 4s), but did have suggestions.

The questions that were most likely to get 4s (4 each) were "Content was helpful", "Content was sufficiently comprehensive", and "The length/format was appropriate". The suggestions also tended to favor content – three people wanted the content to be more basic, including things like the step-by-step process to get files from EUSWeb, and three wanted a separate, advanced training to focus on something different (new process 1, estimates review 2). Three people mentioned projections specifically, either saying that was their primary focus and they wanted more about the integrations of the two or that they wanted specific Projections Suite training.

Two people wanted training to occur more frequently.

There were some comments about facilities – a larger room, having computers with the software already set up, the zoom on the screen could have been larger. Everyone who answered the question said they'd recommend the training to others and a few added positive comments about hospitality or welcome.

After the training we got some feedback from LMI directors that they didn't hear about the training until after it had occurred and would certainly have sent people had they known. The training was promoted first to LEWIS users and filled up very quickly so directors likely would not have heard about it independently.

We have since been approved for online training and we're pursuing that in the near term, but also hope to host more in-person trainings. We have considered the possibility of combining them with other events like the PMP conference, but at the moment that seems unlikely because of the need to host them within a state or federal network for security reasons.

Python Users Group

In 2020 we attempted a Python Users group for LMI users. While interesting, it floundered because we had relatively few participants. In 2023 we reached out to the original participants and others who had expressed interest and have gone through two meetings with some success. There is a GitHub repository for sharing resources and participants have the option of making connections with people who are doing similar work. This will continue, but we've also started conversations about collaborating with NASWA who have considered a similar activity.

WID 3.0

The Analyst Resource Center (ARC) has developed a WID 3.0. A departure from the 2.X versions means larger changes to the database structure, especially to the primary key structure. In regular updates, there are fields added to tables, removed from tables, field type expansions (longer text fields, larger numeric types), table deprecations and additions. The core lookup tables are kept the same and field names are unrevised to minimize impact on dependent applications and automation. By contrast, a major version release will change primary key structures, which will affect both lookup tables and data tables. Some core tables may be dropped, and others designated as core. All tables deprecated in the 2.x versions will be deleted.

However, the current structure of the database has begun to create problems as the technology needs of states are changing, and more and more data are becoming publicly available. Looking forward, there's a need to accommodate those changes with a more adaptable database structure. While many tables will remain unchanged, there are some significant improvements we're considering.

- Adding a version number to the areatype concept. Since the WID was first set up the Office of Management and Budget (OMB) definitions of Metropolitan Statistical Areas (MSAs) has changed from an infrequent occurrence to one that's revisited every few years. As a result, accommodating the new MSA definitions is resulting in a lot of new areatypes and they can easily get out of sync as not every state has changes in every new version. By creating areatype categories and vintages we hope to address those problems.

- A minor revision to the time period lookup tables, so that there isn't the redundancy that we now have. Note that this will not affect data table structure.

- Changing field names. When the WID was first created, FoxPro was a major database option, and field names were restricted in length to accommodate its standards. As a result, many of our field names are cryptic or misleading or inconsistent. As we move forward and length isn't as much of a factor in the software states are now using, giving tables and fields more human-readable names will make them more intuitive and make it easier for new users to understand the structure.

- More significant revisions to some of the non-core data and lookup tables will also be considered.

Feedback

Our policies for making a major version release are:

1. Structure Committee develops the new database structure.
2. Approved by the ARC Consortium present.
3. Given to ARC member states for comments - 30 days.
4. Revise based on ARC comments.
5. Draft of proposed structure is delivered to states and ETA for review and comment.
6. 90 days for initial comments.
7. Structure Committee reviews all comments after 90-day period; responds as necessary and sends out revised draft of structure.
8. Revised draft sent to all states and ETA for comments - 30-day period.
9. Structure Committee reviews second round of comments and makes final revisions to new structure.
10. Final review by ARC member states.
11. Final structure released, along with supporting documents.
12.

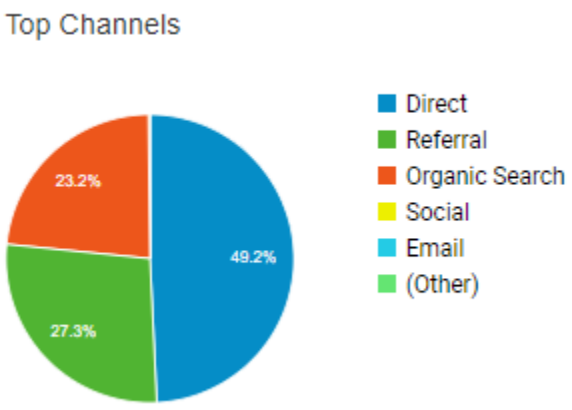
Communications

Website

Google Analytics

We use Google Analytics to keep tabs on what pages users are going to on the site itself. The following refer to recent calendar years.

Half of users are coming directly to our site – users have us bookmarked, come from an email, or type in the address. While a significant portion come from searches, many of those are isolated to a couple of topics and often don't stay long.



Referrals are mostly known partner organizations, primarily CareerOneStop by a large margin. Recently ONetOnline.org has begun sending referrals to the WIDCenter.org website. These appear to be listed as the source for licenses in occupational pages. Ulmita.org and lewissupport.com are both pages associated with the LEWIS software. Links from FloridaJobs.org are declining.

Referrals	2022	2021	2020
careeronestop.org	1,894	2,289	417
onetonline.org	443	-	-
lewissupport.com	32	29	14
ulmita.org	30	10	-
statics.teams.cdn.office.net	18	-	-
floridajobs.org	12	35	46
cn.bing.com	11	7	-
cloudfront.careeronestop.org	8	-	-

Organizational pages describing our tasks and supporting documentation like the TEGL are also a significant driver of traffic. The top link under Technical Documentation is the page ONet links to. The others are reference and documentation and have been largely steady the last few years. Some of the crosswalks have been declining, either because SOC 2018 has now been fully implemented and it's no longer a hot topic or because of changes in the BLS website that linked to legacy crosswalks.

Selected list of high traffic landing pages

	2022		2021		2020	
	pageviews	*unique pageviews*	*pageviews*	*unique pageviews*	*pageviews*	*unique pageviews*
Technical Documentation						
/document/license/	3,522	2,811	3,543	2,672	857	603
/document/all-core-tables/	983	707	997	659	807	517
/wid-downloads/	978	682	1,025	618	944	542
/structure-2/	756	627	711	500	526	400
/document/occcodes/	419	287	468	271	467	277
/workforce-information-database/	410	373	455	364	439	365
/tableau-wdcs/	405	351	732	586	796	586
/arc/update-calendar/	333	260	277	220	272	195
/document/indcodes/	314	214	166	116	147	83
/supported-activities/licenses/	308	249	248	196	211	158
/document/industry/	254	194	214	153	192	130
/document_category/data-tables/	231	157	280	167	221	161
/newsletters-reports/	211	182	194	156	288	208
/document/ces/	200	166	205	132	183	122
/document/labforce/	191	162	258	186	201	139
Organizational Documentation						
/employer-database/	798	671	740	584	482	369
/im-new/	750	597	768	600	671	516
/arc/	592	490	507	436	422	346
/state-dbas/	478	394	495	379	394	315

/tegl/	374	308	399	337	313	255
/why-a-standard-structure/	306	267	317	290	297	266
/training/	304	244	249	197	175	132
/upgrade-the-wid/	169	143	230	160	226	171
Crosswalks						
/document/legacy-crosswalks/	474	294	571	321	441	241
/document/cip-soc-crosswalks/	375	295	614	465	893	616
/document_category/crosswalk-tables/	280	224	447	306	508	346

One of the biggest services we provide to our users is to provide context and references for where to obtain useful data. That section of the site is a major driver of traffic. The largest landing page is about occupational licenses, which is the only data product for which we are the original source. It's also the one that CareerOneStop uses and is the destination for ONET referrals. Core tables are a critical reference for State LMI offices implementing the WID. Legacy crosswalks are older content referring to no longer active taxonomies. This is used by researchers and occasionally linked to by federal agencies.

There have been gains in Indcodes (which didn't make the top 10 in 2021), likely because of the implementation of NAICS 2022. There have been declines in CIP-SOC-Crosswalks, likely because we're past the implementation of SOC 2018.

Content Drilldown	**2022**	**2021**	**2020**
/license/	3,524	3,544	857
/all-core-tables/	983	997	807
/legacy-crosswalks/	474	571	441
/occcodes/	419	468	467
/cip-soc-crosswalks/	375	614	896
/indcodes/	315	-	147
/industry/	254	214	192
/ces/	200	205	183
/labforce/	192	258	201
/education/	152	205	-

Newsletter/Email list

We have an email list with approximately 100 recipients that's mostly used to alert users to new file availability. We also distribute an annual or twice-annual newsletter to that list. In 2020 we converted this from a manually managed list in Outlook to GovDelivery. While users can't add themselves to the mailing list, they can unsubscribe without reaching out. There's also tracking attached to GovDelivery, allowing us to determine which links are popular and how many readers actually open it.

This year only one winter newsletter has gone out. In that, 97% were delivered, with 39% opened. Of recipients, 20% clicked further within the newsletter.

Contact list updates

The technical contact list that the data validators began in 2020 has been running consistently since then, with regular updates to contacts.

Projects

LEWIS

The LEWIS application was originally developed in North Carolina as a means to aggregate Occupational Employment Statistics (OES) survey data to non-published regions while still following the very specific procedures and suppressions defined by the BLS. Because OES data is only published for states and MSAs, many states saw the value in being able to produce reliable estimates for other regions. It also allows wage and occupation filters – some states may use it to produce estimates about minimum wage jobs or combined clusters of occupations. The management of the software was eventually moved to Utah to ensure continuity as its original creator approached retirement and the funding is now managed by the ARC. In 2019/2020 we received approval to begin a new cloud-based version of the application. The original desktop application had many problems – installation in different state IT environments and supporting those differences was cumbersome, security concerns as secure data was stored on a local computer or passwords were shared, and the web-based application has been eagerly anticipated for its easier management.

The transition to the cloud environment and MB3 methodology is now complete and the team's focus is on enhancing estimates review features and providing training to bring users up to speed. One in-person training has been held (described above) and future trainings are being planned.

License data

The Analyst Resource Center (ARC) has been collecting occupational licenses from states since 1997. Historically these were compiled in a central location but have only been made available to the public for several years. The publication is handled by CareerOneStop. Download files are made available on the server to facilitate communications with states but they're not advertised on the main site, except in the license guidance.

While we collect some data centrally and validate state submissions against other data sources that mean that some parts of the data are always more current than this, the following summarizes the most recent submission we got from each state.

Data Release	**States Included**
2022	Hawaii, Tennessee, Ohio, Maine, North Carolina, South Carolina, Washington, Utah, Rhode Island, Oregon, Oklahoma, Mississippi, Maryland, Kentucky, Kansas, Arkansas, Illinois, Colorado, California, North Dakota, Pennsylvania, Georgia
2021	Wyoming, Texas, Vermont, South Dakota, Alabama, Minnesota, Indiana, Louisiana, Nebraska, New York, New Mexico, Wisconsin, Connecticut
2020	New Hampshire, Guam, New Jersey, Michigan, Florida
Earlier	Missouri, Montana, Alaska, Massachusetts, Idaho, Nevada, Virginia, Arizona, West Virginia

Licensed Occupation Admin Tool

To further improve occupational licenses, the next step will need to be the improvement of file transfer and data sharing. Right now a lot of the burden of developing a process to keep licenses updated is

placed on individual states. They're expected to do a lot of background prep work and then submit it in a database structure. The skill sets required – designing a process, collecting, editing and reviewing data, then structuring it in a normalized format – are rarely combined into a single position, so many people are pushed outside their comfort zone for this task. While many states have a very good process in place and any tool would allow them to continue that by just submitting the standard file structure, building some of the review process into an application would reduce the burden on LMI shops that are still establishing that process. This will also make the additions we put in centrally more visible to states and give states opportunities to make changes outside the normal two-year review cycle.

We're currently exploring the possibility of an occupational license administrative tool. We've discussed requirements with Utah and internally and are doing outreach to states to see what features would be helpful. It is likely that we will be able to improve file transfer at the very least.

Employer Database

Job seekers and economic developers often have a need for a list of businesses by region or industry as a contact list. LMI data is protected by CIPSEA and no non-aggregated individually identifiable data can be shared even when the details are no more than what could be found in a phone book. To fill this need there are a number of private companies that offer a range of products, often marketed to libraries and firms building client sales lists. The products are expensive and their off-the-shelf options may not suit state purposes so a joint effort to procure this product for all states was initiated to negotiate better rates, specific deliverables, and to enable even states with more limited economic development budgets to have access.

The contract has been run through ARC for 24 years and in that time there have always been challenges around state procurement. Individual states have been contract holders, and all have been sued by the company that did not receive the contract, creating a costly legal process. In early 2017 ARC approached ETA to procure this through the GSA schedule. The RFP process was allowed to lapse and there were challenges along the way, but the contract with InfoGroup (now DataAxle) was executed on April 3, 2020.

There are a number of additional variables that DataAxle has described and offered either as part of our current contract or at an additional cost. They have provided samples of the data and we have considered their applicability to our users.

Occupational License Holders

The sample data provided was a list of individual license holders, grouped by license category. The coverage was incomplete, it was major license types only. This means that emerging licenses or odd state licenses that we have weak coverage of in our dataset would not be enhanced by the available data. Additionally, the license types were grouped more broadly than standard SOC codes and there was no detail on the requirements or name of the license itself. The quality seemed good – when compared to license types where license holders are publicly listed the majority were direct matches, although the address associated with the license holder appeared to be different (probably associated with an employer rather than home address). There was some potential for research into cross-state licensure, but for the most part the aim of the occupational license data through DataAxle was different than our mandate.

Minority, Veteran, and Woman-owned

This is a topic that DataAxle says will be added as a flag to our current extract. While we did evaluate two samples, both were limited to the firms with a minority, veteran, or woman owner flag. The core use case for this type of data is what percentage of businesses are minority/veteran/or women-owned, but without a total to compare to it was impossible to determine what percentage that would be. From conversations with DataAxle staff, it seems that these attributes are derived from federal programs and are likely accurate in a way we can have high confidence in a positive indicator, but their completeness was more questionable.

QCEW Microdata

Another conversation we participated in was the use of the QCEW address file as a sampling frame. This was a conversation initiated by ONET, who pays a different provider for an appropriately sampled extract of businesses. While their use case was in line with existing uses of the QCEW data (since it's used privately), ours is more public facing and less likely to be approved, in which case the DataAxle contract would continue.

Exploratory Projects

The ARC is in a unique position of needing to know what the next priority of states will be to design database structures and documentation to support those efforts. While our members' connections to other advisory bodies and programs can help track those future efforts, there's also a need to follow topics that are outside government and occasionally experiment with new offerings.

Real-Time analytics

Real-Time analytics of hiring has been a topic of interest to states for several years now. Generally, the term applies to tools provided by private companies that attempt to describe hiring practices based on the information available on the internet. These are either from companies that own websites that are job seeker destinations and host a large number of advertisements on their own, or companies that scrape a pool of major employers to develop metrics. Because of the nature of the source data, the sampling methods aren't traditional and among industry experts the statistical value of real-time analytics based off job postings is debated. Despite this, states remain interested in the topic because they fill a niche that our existing data sources can't. While a few states have a Job Vacancy Survey to look at hiring, most do not and even those that do have infrequent surveys. Real-time analytics are more current than any of our employment statistics – they can be up-to-date to the day, compared to monthly for LAUS and CES. They also are not subject to suppression in the same way and their cost/sample is much lower. At least in theory, this means it's possible to get much more granular data, particularly in rural areas where Labor Market data is scantly available.

Unfortunately, evaluating the quality of the products available has been a challenge. They're expensive, sometimes report as percentages or indexes that don't compare easily to other sources, and their methods are deliberately obscured to protect the companies' intellectual property.

We have been active participants in the NLX Research Hub project put out by NASWA, offering suggestions for how they can improve their APIs. Our exploration of the use of that data in a JVS Survey continue.

Future Directions

There are a number of topics that come up in questions or when designing applications that intersect with our strengths and could be pursued.