

Using Real-Time LMI To Explore Green Jobs

Elizabeth Laird, Brittany English, and Jillian Berk

In February 2009, President Obama signed the American Recovery and Reinvestment Act (Recovery Act) into law to address the employment challenges facing America's workforce. Among other investments, the Recovery Act included \$500 million to support jobs (also known as "green jobs") in the energy-efficiency and renewable-energy industries. While 90 percent of the U.S. Department of Labor's (DOL) Recovery Act green-jobs funding went to support training programs for workers, \$50 million was reserved for grants to state workforce agencies for improving labor market information (LMI). As part of their efforts, 16 of the 30 LMI grantees used real-time LMI to explore employer demand for green workers, the characteristics of green jobs, and the skills required (Laird et al. 2012).

In this brief, we provide some lessons learned from the grantees' experiences using real-time LMI. While the brief draws on experiences and lessons from all of the LMI grantees, it relies mostly on the work of the Northeast Consortium. This consortium, consisting of eight states,¹ dedicated a significant portion of its \$4 million grant to examining the quality of real-time LMI data and the potential of using real-time LMI to make labor demand projections.

What Is Real-Time LMI?

Real-time LMI analysis uses web-scraped job postings to make inferences about the labor market. Many of the grantees also used LMI grant funds to add such job postings to the state labor exchanges; in this brief, however, we focus on attempts to use aggregations of online job postings to provide information on current labor market conditions.

Real-time LMI is collected from a variety of online job posting sources. Vendors have developed "spidering" or "web-scraping" technology that enables them to gather this information

About This Series

In December 2009, the U.S. Department of Labor (DOL) awarded State Labor Market Information (LMI) Improvement grants (LMI grants) to 24 individual state workforce agencies (SWAs) and six consortia of SWAs. Grantees used these LMI grants, which ranged from approximately \$750,000 to \$4 million, to collect, analyze, and disseminate LMI and enhance the labor-exchange infrastructure for jobs and careers within the energy-efficiency and renewable-energy industries. In September 2010, DOL contracted with Mathematica Policy Research to evaluate the extent to which the LMI grant program had achieved its stated purpose. Mathematica was asked to broadly document the activities of all 30 grantees, provide a detailed description of the activities and partnerships of a subset of grantees, and identify grantees' challenges and promising practices. This brief is part of a series that explores lessons from the LMI grant program.

"Real-time LMI is labor market intelligence derived from the analysis of job postings and resumes placed into public and private labor exchanges. It is real time because it can be based on data pulled from the Internet on a daily basis. It is labor market intelligence because it can provide indications of supply and demand trends, emerging occupations, current and emerging skill requirements, and market-based demand for education and certifications."

(Vollman 2010).

¹ Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island, and Vermont.

from Internet websites that host job postings, such as online job boards, employer sites, newspapers, and government websites. After job postings are collected, software parses each job listing into data fields that can be analyzed. Information included in online postings varies; some postings show detailed information, such as required skills, education levels and certifications, wages, and geographic locations, while others display only limited information.

Real-time LMI aims to address the same questions as traditional LMI. Traditional LMI includes data systematically collected from employer surveys, administrative sources such as wage records, and interviews with employers. While real-time LMI and traditional LMI address many of the same topics, there are significant differences in the underlying data sources and the data collection methods. This brief explores some of the strengths and limitations inherent in using web-scraped online job postings as a source of LMI.

Additional Information on Real-Time LMI

To Learn More about Real-Time LMI, Please See

Northeast Consortium's Guide to Using Real-Time Data for LMI Analysts

Jobs for the Future's Aligning Community Colleges to their Local Labor Markets

Brookings Institution's Labor Market Information Forum on Real Time Labor Market Information

To Learn More about the Evaluation of the LMI Grants, Please See

Investing in Labor Market Information (LMI): A Summary of the State LMI Improvement Grants

Investing in Labor Market Information (LMI): Insights from the Recovery Act LMI Grants

How Does Real-Time LMI Work in Practice?

We use two job postings, accessed using two online job boards, to demonstrate data typically captured by spidering or web-scraping, as well as their limitations.

Information Accessible Through Postings

As demonstrated in the green-jobs posting for a “Wind Plant Technician, Leaning Juniper,” (Box 1) LMI analysts can gain an understanding of skills and certifications, firm traits, emerging job titles, and education requirements from an analysis of online postings.

Skills and certifications. This posting includes detailed information about the skills and qualifications necessary for this position. However, a scraping or spidering program may have difficulty in distinguishing between the listed skills, such as “advanced computer skills utilizing word processing, spreadsheets, email, facility control, management and reporting systems...” and qualifications, such as a candidate having a technical certificate instead of college experience.

Firm traits. When job postings such as this one clearly identify the firm's name, the information allows LMI analysts to conduct firm-based analysis, which includes examining what types of jobs the firm posts and how frequently it posts openings online. Analysts may also be able to use the firm name to link real-time information with existing administrative or survey firm data.

Frequency of job titles. In addition to detailing emerging skills, job postings allow policymakers to examine the frequency of postings for particular job titles. For example, by using job postings, analysts can determine if the number of openings for such jobs as “Wind Plant Technician” are increasing or decreasing.

Education requirements. Not all job postings include education requirements, since degrees may be implied by the job title, as is the case for lawyers and doctors. For other positions, job postings may include preferred or required education levels. For example, this firm requires that applicants have “one year college or a technical school certificate; or 12 months of related experience or wind industry training; or equivalent combination of education and experience.” Although the posting includes detailed education information, parsing this posting into data fields that can be aggregated and analyzed may be difficult.

Box 1. Job Posting for Wind Technician

Wind Plant Technician, Leaning Juniper-120042**Description**

██████████ is seeking a Wind Plant Technician for the Leaning Juniper facility located in Arlington, Oregon. This position is responsible for the onsite operations, maintenance, repairs and replacement of equipment on a wind powered generation plant, and leads in all areas of plant operations as directed by the Plant Manager.

Responsibilities Include:

- Operate and maintain wind turbines and associated wind plant equipment, including substation and tower switchgear.
- Read schematics to troubleshoot complicated mechanical, electrical, and hydraulic problems with variable pitch, variable speed control systems and components.
- Perform mechanical and electrical component repair or replacement of parts to correct malfunctions following all manufacturers' requirements.
- Comply with all project environmental health and safety programs.
- Perform preventive maintenance in accordance with OEM maintenance manuals.
- Adhere to effective internal controls.

Qualifications

- One year college or a technical school certificate; or 12 months of related experience or wind industry training; or equivalent combination of education and experience.
- Valid driver's license and acceptable driving record.
- Ability to climb 80 meter steel towers on a daily basis to perform maintenance in the nacelle and hub.
- Ability to frequently lift items weighing 1 to 60 pounds and occasionally lift items from 61 to 100 pounds.
- Advanced computer skills utilizing word processing, spreadsheets, email, facility control, management and reporting systems, including the use of computerized diagnostic and troubleshooting tools.
- Ability to interpret and follow a variety of instructions furnished in written, oral, diagram, or schedule form.
- Possession of comprehensive writing skills to meet contractual reporting obligations.
- Willing to work overtime on holidays, weekends, and on short notice. Participates in an on-call schedule.
- Ability to work in extremely adverse weather conditions.
- Ability to work around low and medium voltage.
- Ability to solve practical problems and deal with a variety of concrete variables in situations where only limited standardization exists.
- Weight must not exceed 285 lbs (excluding equipment/gear) when performing job related tasks.
- Knowledge of heavy equipment required for facility maintenance with a working knowledge of crane and rigging requirements.
- Familiarity with maintenance and electronic testing equipment. (i.e. voltage testers, amp clamps, oscilloscopes, megohmmeters, infrared testers, hydraulic torque equipment, alignment and various hand and power tools.)
- Willingness and desire to increase professional knowledge of turbine equipment, including safety regulations.
- Willing to travel to other wind sites within U.S. as required.
- If selected for an interview, candidates will be asked to review the Functional Job Analysis requirements for this job.

Preference may be given to candidates with the following:

- Experience in power plant operations and maintenance.
- Experience in wind plant operations and turbine experience, including specific experience in wind turbine power converters.
- Experience in high voltage switching.
- Qualified candidates living in or near the greater local area.

All offers of employment are contingent upon the successful completion of a medical Fit for Duty exam, background and reference check, drug screen, verification of legal right to work in the U.S., and in some cases, a credit check. A credit check will be administered when a prospective employee will be working in Finance, Accounting, Treasury or where duties may involve handling of funds, accounts or cash. A Motor Vehicle (MVR) check will be administered when a prospective employee will be regularly using a company vehicle. Iberdrola Renewables is an Equal Opportunity Employer.

Job Category Craft/Skilled/Technicians**Primary Location** OR-Arlington

Box 2. Job Posting for Installation Installer

Job Summary

Company
[REDACTED] INSULATION

Location
Portland, OR 97223

Industries

- Construction - Industrial Facilities and Infrastructure
- Construction - Residential & Commercial/Office

Job Type

- Full Time
- Employee

Years of Experience
2+ to 5 Years

Education Level
High School or equivalent

Career Level
Experienced (Non-Manager)

INSULATION INSTALLER

About the Job

Insulation Contractor in Portland Oregon seeks Experienced Insulation Installers

Insulation contractor seeking insulation installers with install experience in new residential and commercial construction, multi-family and weatherization projects. Installers must have previous employment with professional insulation contractors.

Good driving record and driver license is required.

Competitive piece-rate and benefits.

Please email resume to:
[REDACTED]

[APPLY](#)

JOB TOOLS

- [Print](#)
- [Save](#)
- [Email](#)
- [Share](#)
- [Report this job](#)

MORE JOBS

[REDACTED]

- [View similar jobs](#)
- [All PORTLAND, OR jobs](#)
- [Browse all jobs](#)

Source: Available at [http://jobview.monster.com/INSULATION-INSTALLER-Job-PORTLAND-OR-US-113362044.aspx]. Accessed August 15, 2012.

Limitations of Real-Time LMI

Although we can locate a good deal of information from “scraping” online job postings, their use for LMI analysis faces several limitations. At its best, real-time LMI can present information on the universe of online job postings. If certain types of jobs are not advertised online, they will be underrepresented in any labor market analysis based on real-time data. Although the share of job listings posted online has been growing, some jobs, including certain low-wage or union positions, may not be captured fully.

Online job postings are designed to meet the needs of employers, not those of LMI analysts. Employers use postings to attract applicants; consequently, postings tend to include vague details so as to attract a large number of applicants. Using a job posting for “Insulation Installers,” (Box 2) as well as the first posting, we explored some of these limitations.

Postings do not equal job vacancies. In the case of the insulation installer posting, it is unclear if the posting represents one vacancy or multiple vacancies. Although this posting seems to represent more than one opening, since it uses the plural in “seeking insulation installers,” many firms also use online job postings to collect resumes for projected rather than actual vacancies, as could be the case for this posting.

Geographic information may be inaccurate. The location for this posting is listed at the city/state level. However, this

geographic information may be for a corporate headquarters rather than the location of the vacancy. For example, in this posting, users know that the firm is located in Portland, Oregon but cannot discern with certainty that the job openings are also in Portland. Given the potential inaccuracies in the geographic information, analysis of aggregate job postings may lead to inaccurate data on local labor market conditions and the geographic locations of job growth.

Salary information typically is excluded. Like many job postings, this one does not include detailed salary information. Instead, it states “competitive piece rate and benefits.” (The wind plant technician posting provides no salary information.) Analysts and job seekers thus may need to use traditional LMI on average occupational salaries to supplement the information available in online postings.

How Does Real-Time LMI Help Us Better Understand Green Jobs?

Real-time LMI can highlight emerging areas of economic interest such as green jobs through its use of current online postings. As part of the LMI grant program, analysts used aggregated online job postings from identified green jobs to assess employer demand, determine required skills, and project job growth.

Identifying Green Jobs

Grantees who wanted to use real-time LMI to better understand green jobs faced significant challenges. To capture real-time LMI for green jobs, one of the first tasks was to define “green.” Most grantees identified online job postings as green based on occupation and industry. The Northeast Consortium took a different approach, attempting to identify jobs based on “green words” in the posting.

Occupation and industry codes. Under this approach, grantees identified jobs as green based on occupation and industry. Some grantees used the list of green occupations developed by O*NET. Others supplemented this list with additional occupations identified through their survey efforts. Identifying green jobs based on occupation and industry has two important limitations. The first is that not all job listings clearly identify the occupation and industry. The second is that this approach may result in significant over-identification of green jobs. Not all positions in a given occupation are green; using this approach, however, all postings under a given code would be flagged. For example, all chemical engineer postings would be identified as green jobs, even though not all chemical engineer positions are green. Conversely, certain green job postings may not be flagged because the occupation or industry is not on the green list.

Parsing tool. An alternative approach, developed by the Northeast Consortium in collaboration with its vendor, involved using a list of green phrases to identify postings as green. The vendor’s software tool parses the text of an online job listing, searching for green phrases. The development of this parsing software and a list of green terms was an iterative process. The Northeast Consortium would examine postings identified as green as well as those not identified as such to determine the success of the parsing tool and refine its procedures and green terms list. By the end of the grant, the list had grown to nearly 900 key phrases. While this approach may accurately identify green jobs, it is time and labor intensive. Developing an accurate list required extensive knowledge of the labor market and a commitment to updating the list as new information becomes available.

Analyzing Green Jobs

Once jobs were identified as green, the LMI grantees used the real-time LMI for multiple purposes. Specifically, LMI grantees assessed the demand for green jobs at a specific point in time; determined skills, certifications, and training requirements for green jobs; and explored using real-time data as a source for short-term projections.

Assess demand. Grantees wanted to use real-time job postings to assess the employer demand for green jobs in their local labor markets. Indiana, a member of the Driving Change consortium, an LMI grantee, used real-time data to assess demand, but acknowledged that the data source was imprecise because it was not always possible to be sure that the posting reflected an actual vacancy. Instead of using real-time postings as an exact measure of job openings, it used real-time data “to gauge the relative strength of current employment demand by calculating the ratio of [real-time] postings to the average 2009 employment” (Driving Change 2011). Grantees also examined job posting data to determine the occupations and industries that place the most ads for green jobs.

Determine associated skills. Real-time data allowed users to track the demand for specific skills and certifications necessary for green jobs. For instance, posted jobs from known green firms that required certifications, such as Leadership in Energy and Environment Design (LEED) certifications, provided insight into what was required for positions in green industries. In its analysis of job postings, the Northeast Consortium determined that the top three green “skills” included in job ads were environmental health, water treatment, and LEED certification.

Project job growth. Despite the challenges in using real-time data as described above, the Northeast Consortium explored using these data as a source for short-term employment projection. However, an additional challenge for doing projections was the need for repeated consistent measures of employment demand that analysts could use to construct a demand model. If a vendor starts to spider additional websites or changes the frequency of the spidering, then the measures are not consistent and job posting data may exhibit a spike that is purely an artifact of the change in data collection parameters. One solution proposed by the Northeast Consortium is to work with real-time vendors to develop a more limited set of websites that are spidered consistently. Since the Northeast Consortium’s real-time vendor could not implement this approach retroactively, the Consortium did not generate employment projections; instead it developed a projections methodology that can be used in the future to generate short-term employment projections.

What Lessons Have We Learned from Real-Time LMI Use?

Analysis of real-time data, particularly by the Northeast Consortium, shows that these data have value as a point-in-time measure to capture information on new and emerging occupations and market-demand skills. Real-time LMI provides an opportunity to analyze current labor market conditions without having to wait for the lag present in traditional LMI. The expe-

periences of the LMI grantees suggest that when using real-time data, the following lessons are important to consider: data quality is central; and real-time LMI provides better information when combined with traditional LMI.

Users should be cautious about data quality. Real-time data can be noisy, producing a large number of false positives from duplicate observations, inaccurate flagging as green, and postings from employers that do not actually have current open positions but want a constant applicant pool. The Northeast Consortium found that constant human monitoring was a necessary prerequisite for high quality data, and that real-time LMI needed the “same level of quality control that is common in BLS sponsored surveys” (Northeast Consortium 2012). Researchers using real-time data need to understand what sites are being spidered and how frequently job postings are being scraped from the sites. Any changes in the underlying web-scraping procedures can produce the appearance of a change in the underlying economic conditions. It is particularly challenging to construct a time series in an environment in which the actual data collection process is constantly evolving.

Real-time LMI provides better information when combined with traditional LMI. Real-time data provide users with a snapshot of the current market and may “prove to be the best source of data on certain credentials (e.g., industry-based certifications) that are not captured in current surveys” (Northeast Consortium 2012). Real-time data also provide information on new and emerging occupations that may take more time to appear in traditional LMI; for the best use of real-time data, however, users should combine real-time analysis with tradi-

tional LMI. This combination provides context and additional information to supplement data gathered from online postings.

References

- Berk, Jillian, Diane Herz, Elizabeth Laird, Megan Hague Angus, Brittany English, and Lauren Bernstein. “Investing in Labor Market Information: Insights from the Recovery Act LMI Grants.” Washington, DC: Mathematica Policy Research, September 2012.
- Driving Change. “Driving Workforce Change: Regional Impact and Implications of the Auto Industry Transformation to a Green Economy.” Available at [<http://www.drivingworkforcechange.org/reports/DrivingWorkforceChange.pdf>]. Accessed August 10, 2012.
- Laird, Elizabeth, Jillian Berk, and Brittany English. “Investing in Labor Market Information: A Summary of the State Labor Market Information Improvement Grants.” Washington, DC: Mathematica Policy Research, August 2012.
- Northeast Consortium. “Understanding the Labor Market in New Ways. Final Report.” Northeast Research Consortium, March 2012. Available at [<http://public.greencarecentral.org/Final%20Report.pdf>]. Accessed August 11, 2012.

This report was prepared for the U.S. Department of Labor (DOL), Employment and Training Administration, Office of Policy Development and Research by Mathematica Policy Research, Inc., under contract number DOLQ091A20941. The views expressed are those of the authors and should not be attributed to DOL, nor does mention of trade names, commercial products, organizations imply endorsement of same by the U.S. Government.

For more information, please contact Jillian Berk at jberk@mathematica-mpr.com.

For further information on this issue brief series, visit our website at www.mathematica-mpr.com

Princeton, NJ • Ann Arbor, MI • Cambridge, MA • Chicago, IL • Oakland, CA • Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research, Inc.