
THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

Overview of Private Sources of Workforce Data

Prepared by the Center for Regional Economic Competitiveness for
GWIPP

March 2015

MEMORANDUM

TO: ANDREW REAMER
FROM: KEN POOLE, BRENDAN BUFF, ANIKA VAN EATON
SUBJECT: LUMINA – NONFEDERAL DATA SOURCES
DATE: MARCH 18, 2015

During the past few months, CREC has gathered data about several private sector data providers that could have some value to the education and workforce system. These data sources offer unique approaches to understanding educational attainment and workforce progress. The data sources were identified based on our experience with education and workforce agencies as well as based on primary research.

The list of vendors, including both for-profit and not-for-profit organizations, is designed to be illustrative of how the private sector is responding to the increased demand for data that could be used by policymakers. The 33 vendors identified below do not represent an exhaustive list because so much innovation and experimentation is occurring in this space.

In examining the vendors and data programs identified in this collection, CREC categorized the programs into five broad types:

1. Vendors with data analytics products already designed for specific aspects of the public education and workforce development market
2. Vendors with products that could be adapted to the public education and workforce development market
3. Vendors with an interest in the public education and workforce development market that are also in the process of developing data/analytics products
4. Vendors that have potentially useful data products but for which their interest and in the public education and workforce development market is unclear
5. Public data aggregators that are likely to incorporate public and/or proprietary education and workforce development-related data in their product

These categories are roughly prioritized based on our assessment of how useful or connected to the national education and/or labor market information (LMI) data system the vendors or their products are, with the first category being the most highly connected already. However, within these groups it is possible to draw further distinctions. These data organizations have different target consumers and types of product offerings. Human resource offices, researchers, higher educational agencies, workforce programs, and state labor market information shops all serve as consumers of these organizations' products. Conversely, product offerings might include IT support, data aggregation, or original data content.

- 1. Data analytics product vendors already marketing to the public education and workforce system**
 - Burning Glass Technologies
 - EMSI/Career Builder
 - Geographic Solutions

- National Student Clearinghouse
- TORQWorks
- WANTED Analytics/Help Wanted OnLine

These companies have all established themselves already with the higher education and LMI system. Burning Glass, Geographic Solutions, and WANTED Analytics are content providers that produce real-time labor market information. EMSI is a data aggregator that began by targeting the community college system and has since gained a strong foothold with workforce agencies. More recently, EMSI has added capabilities to provide web-scraped job advertisements, making them a competitor as “real-time LMI” data providers. The National Student Clearinghouse is a content provider for the higher education system, with a warehouse of student data including enrollment and degree records that could potentially be an invaluable resource for the public workforce system as well.

The second through fifth categories of data providers represent companies that could potentially contribute data to the national education and LMI system.

Most prominently, the second and third categories include organizations such as the National Labor Exchange, Economic Research Institute, Glassdoor, and LinkedIn that provide existing products that we judged had the greatest potential to be adapted as relevant data to policymakers analyzing educational and labor market outcomes. Our assessment revealed that each of these companies either has a data product in search of a market or have data content that could become useful with effective analytics.

2. Private resource providers that are exploring new markets that could include data users

- ACT
- CareerBuilder
- Economic Research Institute
- National Labor Exchange

In particular, ACT, CareerBuilder, Economic Research Institute, and the National Labor Exchange have already positioned themselves within the education and workforce development field to provide relevant services, but perhaps not direct education or labor market products. These existing products, such as applicant assessments, job and resume posting services, and regional salary comparisons, could easily be translated into useful products. Furthermore, these companies already target job seekers and employers as customers. Data from their existing products could guide decision making about career opportunities or about the labor market supply.

ACT, as a content provider for the education and workforce system, could be a source for data about skills gaps, occupations, competencies, and credentials. CareerBuilder could be a source for data about job postings and skills on resumes. They, along with the Economic Research Institute, are content providers for human resources departments in companies nationwide. The National Labor Exchange—an IT support product for workforce systems managed by the National Association of State Workforce Agencies in collaboration with numerous private employers—provides job postings that could be translated into a source for employment trends, job postings, skills on resumes, and salaries.

3. Resources that have expressed interest in and are developing data/analytics products for the public education and workforce market

- Glassdoor
- Indeed
- LinkedIn

Glassdoor, Indeed, and LinkedIn are all vendors that have services designed for the education and workforce development market, and they also seem to be in the process of developing analytics. They have made explicit pivots toward creating new products to serve public education and workforce development. For instance, in 2014, LinkedIn worked with the City of New York to identify skills gaps to train and hire New Yorkers for technology jobs. LinkedIn is working with the White House’s TechHire initiative to use their Economic Graph to help interested communities assist job seekers decide where to focus their skills development and highlight the most in-demand jobs. LinkedIn is also supporting the TechHire initiative to find better ways to connect people to training and to connect job seekers to employers. Meanwhile, Glassdoor is exploring analytics to use their interview question, salary, and company review data.

Glassdoor and LinkedIn could potentially supply information about trends in supply of and demand for different skills. They are content providers for the workforce system and HR departments, respectively. Glassdoor has data from interviews, which could show trends in the most sought after qualifications. LinkedIn profiles are user-supplied repositories of online badges, certifications, groups, and other signalers. They are also sources of job history, professional connections, and post job openings. These resources could contribute to adding nuance to the skills and credentials of the job market. One challenge for Glassdoor data is that many companies have small sample sizes of employees who felt anonymous enough or comfortable enough sharing their experiences. The drawback to LinkedIn data is that profiles likely do not represent large swaths of the American workforce.

An interview with Indeed revealed that the company has hired a chief economist to help improve the quality of the company’s archived data so that it can begin to provide more systematic analytics about the job postings that the company aggregates. A key challenge that companies like Indeed must overcome in developing data analytics is enhancing the quality of the data collected and ensuring that the information collected is internally consistent and historically maintained, a major data “hygiene” challenge that seldom provides immediate return for information technology companies selling other types of products or services. Thus, job postings must be more standardized; resumes must use more consistent terminology for skills and jobs; and computer learning must account for data inputters who are always as accurate as most researchers would want.

4. Private resource providers in which the extent of data product or market involvement remains unclear

- Certification Registry
- ConnectCubed
- Cornerstone OnDemand
- getTalent
- Gild
- IBM Smarter Workforce
- Infor Talent Source
- Knack
- PayScale
- Simply Hired
- Sociometric Solutions
- Taleo
- Ultimate Software: UltiPro Human Capital Management
- The Vault

George Washington Institute of Public Policy research professors, in partnership with the Lumina Foundation, are in the process of developing the Certification Registry. This Registry will likely

supply useful data on non-degree certifications and badges once it has been launched by developing a common language for the emerging array of certifications and building a database to house this information. In particular, it could be used as a data aggregator for workforce systems or HR departments interested in workforce credentials and what credentials employers require of their prospective employees.

There has been a rise of talent management companies and start-ups that are attempting to streamline the recruiting, hiring, and onboarding processes. Some companies, such as Knack, are attempting to build new ways of testing the skills of job applicants. Knack is also working with the White House TechHire initiative to offer its predictive analytics for IT jobs free of charge to employers, communities, and training providers aiming to identify IT talent among people of color, women, and veterans. As Knack grows, it could become a source of data for the skills of the workforce.

Cornerstone OnDemand, getTalent, Infor Talent Science, and Ultimate Software are talent management companies that streamline recruiting, hiring, and onboarding. While these companies likely have significant data on skills and salaries, it is less likely that they could play a role in the national LMI system because they do not appear to be focused on developing analytics as a spin-off of their products. Furthermore, much of the data they generate fulfill clients' human resources function, raising serious privacy and legality concerns with respect to sharing it. Another company, Sociometric Solutions, focuses more on identifying competencies and evaluating team performance for human resource departments, offering the potential for aggregated data that might provide insights about the prevalence of skills in the job applicant pool or within the incumbent workforce.

User-driven websites such as PayScale and Simply Hired could come to serve a similar role as Glassdoor and other job board companies. PayScale has significant data on salaries and already provides analytics to universities to assess how their alumni are faring in the job market. Simply Hired has the job board and company information that could make it helpful in the same way that Career Builder, Monster, Indeed or other job boards could be. Data analytics from these large, anonymized data sets have more potential for contributing to the national LMI system; however, the greatest challenge continues to be finding how best to benchmark these datasets to understand their respective biases in representing the universe of available jobs and available job seekers.

5. Data aggregators that could possibly integrate education and workforce data into relevant IT tools

- Amazon web services
- Google public data
- IBM ManyEyes
- NORC at the University of Chicago
- Open Science Data Cloud
- Quandl
- Statista

There are a number of companies that provide more general data aggregation services that could be applied to education and workforce development activities. Several of these companies aggregate publicly available data sets. These aggregation services sometimes incorporate education and workforce data, but frequently they have accessed a variety of public statistics and administrative records. Available on a subscription basis, some also analyze data and create visualizations for clients. These data services seem like a middle man for accessing useful data, and they seem less useful for contributing to the national LMI system.

Overview of Nonfederal Data Sources for Lumina Project
March 6, 2015

Company	Contact	Key Products	Data Content	Value to Public ED/WF system	Required job skills	Held job skills	Job postings	Compensation
Data Analytics Product and Market in Place								
Burning Glass Technologies	Steve Lynch	<ul style="list-style-type: none"> • Labor/Insight™ • Real-Time Jobs Feed • Applied Research 	<ul style="list-style-type: none"> • Job postings • Skills and qualifications • Locations 	Source for aggregated real-time LMI, including job postings.	x		x	
EMSI / Career Builder	Rob Sikritz	<ul style="list-style-type: none"> • Analyst: • Career Coach: • Professional Services: 	<ul style="list-style-type: none"> • Job postings • LMI 	Analytics targeted towards workforce, economic development, and higher education.	x		x	
Geographic Solutions	Paul Toomey	<ul style="list-style-type: none"> • Virtual OneStop • Virtual LMI • Reemployment Exchange (REX) • Geographic Solutions Unemployment System (GUS) • America's Labor Market Analyzer (ALMA) • America's Virtual OneStop (AVOS) • Labor Market Survey (VOSurvey) • Job Aggregation (VOSJobs) 	<ul style="list-style-type: none"> • Job postings • Unemployment data • LMI 	Provides job postings and regional labor market information.	x		x	
National Student Clearinghouse	Don Hossler	<ul style="list-style-type: none"> • Signature Reports • Snapshot Reports • Current Term Enrollment Estimates 	<ul style="list-style-type: none"> • Enrollment data • Degree records • Demographic data 	Provides student data.	x			
TORQWorks		<ul style="list-style-type: none"> • Career transition tools • Online tool for One-Stops to guide workers to new jobs • Maps non-military occupations to civilians and suggests alternative occupations • Evaluates occupational transferability through the TORQ algorithm. Based on O*NET, state LMI data, and other sources. 	<ul style="list-style-type: none"> • Occupation data • Real time job availability • Training programs • Knowledge, skills, and abilities gaps 	Analysis of related occupations and on similar skill requirements.	x			

Wanted Analytics/Help Wanted OnLine	Bob Plumber/Carol Courter	<ul style="list-style-type: none"> • Hiring Scale • Global Data • Historical Hiring • Posting Period • Search Criteria • Hiring Demand Dashboard • Candidate Rolodex • Heat Map • Market Salary • Skills & Certifications • Graduates Programs • Job Listing • Candidates Supply 	<ul style="list-style-type: none"> • Job postings • Skills and qualifications • Locations • Job seekers • Salaries 	Source for job postings, skills, competencies, and salaries.	x	x	x	x
Product in place; seeking new market								
ACT		Workforce assessments <ul style="list-style-type: none"> • National Career Readiness Certificate • Job Analysis Tools • Workplace Assessments • Skill Gap Training • Evidence-based Credentials • Certificate Reporting & Management System 	<ul style="list-style-type: none"> • Skills • Occupation data • Knowledge, skills, and abilities gaps 	Source for data about skills gaps, occupations, competencies, and credentials.	x			
CareerBuilder		<ul style="list-style-type: none"> • Careerbuilder1 • Broadbean • Workforce Data 	<ul style="list-style-type: none"> • Job postings • Resume postings 	Source for job postings and resume data.	x	x	x	
Economic Research Institute		Compensation comparisons <ul style="list-style-type: none"> • Salary assessor • Executive compensation assessor • Geographic assessor • Relocation assessor • Nonprofit executive position compensation assessor • Occupational assessor 	<ul style="list-style-type: none"> • Salary survey data for 6,000 positions in more than 1,000 industries and over 8,000 locations • Compensation data for over 500 top management positions. • Create geographic salary differentials for more than 8,000 locations in the US, Canada and Europe. • Compare cost of living in over 10,000 locations in more than 200 countries • Determine competitive pay for 90 nonprofit executive positions. Historic data for over 200,000 organizations. • Perform disability determination analyses and overtime exemption testing of over 19,000 position titles using FLSA and 9 state level resources. 	Cross-regional salary comparisons.				x

Indeed	Tara Sinclair	<ul style="list-style-type: none"> • Post a job • Find Resumes • Search Jobs • Analyses • Resume database • Queries trends • Job trends 	<ul style="list-style-type: none"> • Job postings • Resume postings • Queries trends • Job trends 	Queries trends data would be useful in finding the most common job searches. Source for job postings and resumes.	x	x	x	
National Labor Exchange		<ul style="list-style-type: none"> • Post Jobs • Search Resumes • CareerOneStop • America's Career InfoNet • Downloads and Uploads of Job Openings • Indexing (a No Cost Customizable Spider) • Jobs Microsites • Hosted Sites • Analytics 	<ul style="list-style-type: none"> • Employment trends • Job postings • Skills and certifications • Salaries 	Source for employment trends, job postings, skills, and salaries.	x		x	x
Interest in market; in process of developing data / analytics product								
Glassdoor	Andrew Chamberlain, Chief Economist	<p>Company and job resource</p> <ul style="list-style-type: none"> • Job postings <p>User generated data on:</p> <ul style="list-style-type: none"> • Company reviews • Interview reviews • Salary information 	<ul style="list-style-type: none"> • Job postings • Company reviews • Interview questions • Salaries 	Data from interviews could show skills in demand.	x		x	x
LinkedIn		<ul style="list-style-type: none"> • LinkedIn Job Postings • Recruiter • Manage pipeline of talent 	<ul style="list-style-type: none"> • Job postings • Resumes 	Source of skills, online badges, certifications, groups or other signalers people will add to their profiles. Source of job history and professional connections. Source of job postings.	x		x	
Sociometric Solutions		Electronic wearable sensing devices to measure social behavior and generate feedback to assist individuals and teams improve behavior and communication.	<p>Sociometric Badges:</p> <ul style="list-style-type: none"> • Capture face-to-face interactions • Extract social signals from speech and body movement • Measure proximity and location of users 	Source for identifying competencies and evaluating team performance. Source for insights on the categorization of soft skills.				
Extent of data product or market involvement is unclear								
Certification Registry	Bob Sheets/Steve Crawford	<ul style="list-style-type: none"> • Certification language standardization • Certification organization catalogue • Certification Registry - Online database 	<ul style="list-style-type: none"> • Credentials • Certifying organizations • Credential transfer value 	Provides a new source of data on a burgeoning field of certifications	x	x		

ConnectCubed	Michael Tanenbaum, Co-founder / CEO	<ul style="list-style-type: none"> • Candidate assessment games • Employee performance benchmarking 	<p>Results from assessments:</p> <ul style="list-style-type: none"> • Personality traits • Reaction times, focus, and mental bandwidth from memory games • Subject matter knowledge • Spatial reasoning skills <p>Data from benchmarking</p> <ul style="list-style-type: none"> • Performance reviews • Job descriptions • Interview scores 	Source for competencies / skills	x	x		
Cornerstone OnDemand	Michael Housman, Chief Analytics Officer	Software to organize recruitment, onboarding, connecting employees, training, performance, compensation, and succession planning.	<ul style="list-style-type: none"> • Applicant information • Employee training completion, certifications, and goals • Employee performance information (skills and competencies, feedback from team and supervisor) • Assessments of employee potential for succession planning 	Source for competencies / skills, learning about training needs and compensation		x		
getTalent	acquired by Dice.com in 2013. Unclear about status of getTalent functions.	<p>Talent database with resume component</p> <ul style="list-style-type: none"> • Creates pipeline of future potential applicants • People can submit resumes when there is not a job open • Parses resumes to create searchable database • Mobile talent acquisition: App for taking pictures of resumes and entering them into searchable database • Parses resumes to create searchable database 	Resume data	Source to see who is applying for jobs and more about their skills and competencies		x		
Gild	Jonathan Foley, VP of Science	<p>Recruitment resources for tech companies seeking computer programmers</p> <ul style="list-style-type: none"> • Assessment of computer programmer's technical ability through open data sources • Assessment of computer programmer's demand in job market 	<p>Information about computer programmers:</p> <ul style="list-style-type: none"> • Coding skills • Education • Work history • Location • Social site profiles • Contact information 	Potential profiles of resumes and competencies specific to computer programmers.		x		

IBM Smarter Workforce		Talent management software <ul style="list-style-type: none"> • Software to for hiring, onboarding, training, virtual employee sharing, feedback and recognition • Candidate assessments • Retention analytics • Retention analytics 	<ul style="list-style-type: none"> • Applicant information • Employee training completion • Employee performance information (skills and competencies, feedback from team and supervisor) 	Source for job skills and competencies.		x		
Infor Talent Science		Talent management software <ul style="list-style-type: none"> • Cloud-based predictive analytics software to improve employee selection, retention, and training. • Employee performance profiles • Candidate management workflow • Candidate assessments • Employee development module 	<ul style="list-style-type: none"> • Applicant information • Employee training completion, certifications, and goals • Employee performance information (skills and competencies, feedback from team and supervisor) 	Source for job skills and competencies.		x		x
Knack	Guy Halfteck, Founder/CEO	Candidate assessment games	Personality traits resulting from candidate assessments	Source for job skills and competencies.		x		
PayScale		Salary and compensation resource <ul style="list-style-type: none"> • Salary assessments, career planning, and cost of living research for individuals • Employer research • Cloud compensation software w/salary dataset and analytics to help businesses set compensation rates 	<ul style="list-style-type: none"> • User-generated salary, employment, education, and location information 	Free salary calculators, pay and career charts, and data visualizations. A variety of data licensing and partnership opportunities to republic Pascale data. Alumni analytics to assess how college alumni are faring in the job market.				x
Simply Hired		<ul style="list-style-type: none"> • Local Jobs • Company Directory • Browse Jobs • Salary Estimator • Post a Job 	<ul style="list-style-type: none"> • Job postings • Company data • Salaries 	Source for job postings, skills in demand, types of jobs in demand, salaries.			x	x
Taleo		Cloud based solution for: <ul style="list-style-type: none"> • Talent management (recruiting, onboarding, career development, performance management, and succession management) • HR • Payroll • Compensation 	<ul style="list-style-type: none"> • Performance of employees • Education and work history of employees and candidates • Salary information 	Source for skills and salaries.		x		x
The Vault		Rankings and job board <ul style="list-style-type: none"> • Company, internship program, and school rankings • Job board • Company reviews 	<ul style="list-style-type: none"> • Job postings • User-generated company reviews 	Source for company reviews and job postings.			x	

Ultimate Software: UltiPro Human Capital Management		Cloud-based solution for: <ul style="list-style-type: none"> • Talent management (recruiting, onboarding, career development, performance management, and succession management) • HR • Payroll 	<ul style="list-style-type: none"> • Performance of employees • Education and work history of employees and candidates • Salary information 	Source for skills and salaries.				x		x
Public data aggregators										
Amazon web services Public Data Set		Data aggregator	Different data sets about wide ranges of topics	Public data with unique user interface, not specific to workforce and education.						
Google Public Data		Data aggregator	Different data sets about wide ranges of topics	Public data with unique user interface, not specific to workforce and education.						
IBM ManyEyes		Data aggregator	Different data sets about wide ranges of topics	Public data with unique user interface, not specific to workforce and education.						
NORC at the University of Chicago		Confidential microdata management	Different data sets about wide ranges of topics	Portal to private data for vetted researchers.						
Open Science Data Cloud		Science data aggregator	Science data	Source for science data.						
Quandl	Abraham Thomas, Founder/CEO/Chief Data Officer	Data aggregator	Different data sets about wide ranges of topics	Public data with unique user interface, not specific to workforce and education.						
Statista		<ul style="list-style-type: none"> • Data aggregator • Research and analysis • Infographics and visualization 	Different data sets about wide ranges of topics	Public data with unique user interface, not specific to workforce and education.						