Program Evaluations using Big Data

Daniel Querejazu
SRI International
Center for Innovation Strategy and Policy
Who We Are

- An independent, nonprofit corporation
  - Founded by Stanford University in 1946 as Stanford Research Institute
  - Independent in 1970
- Revenues: approximately $540 million
- About 2,000 employees
- 21 locations worldwide
Our Point of View: Bridging Basic Research & Commercialization
About the Center for Innovation Strategy and Policy

The Center for Innovation Strategy and Policy (CISP) helps clients achieve long-term economic and social impact by through effective investments in science, technology, and innovation.

- Build Innovation Systems
- Evaluate Complex Programs
- Transition Technology to Market
- Advance Innovation in Government
About the Center for Innovation Strategy and Policy

The Center for Innovation Strategy and Policy (CISP) helps clients achieve long-term economic and social impact by through effective investments in science, technology, and innovation.

- Build Innovation Systems
- Evaluate Complex Programs
- Transition Technology to Market
- Advance Innovation in Government
Evaluating Investments in Human Capital in the Dept. of Energy’s Industrial Assessment Center Program

Christina Freyman
Roland Stephen
Daniel Querejazu

Center for Innovation Strategy and Policy
Overview

What we did:

• A mixed-methods evaluation design to measure the program’s contribution to training the next generation of energy efficiency engineers.

How we did it:

• Qualitative data from interviews with stakeholders, a survey of IAC alumni, the IAC students exit survey, and SRI’s novel text-analytics-based approach that compared IAC alumni resumes with two comparison groups.

What I will concentrate on here:

• Combined both traditional approaches (interviews and a survey) with new approaches in text analytics (based on unique data sources) to develop reliable and useful measures of the impacts of the program on human capital
Program Overview and Assessment Design
Overview
Program Goals

Efficiency

• Increase the energy efficiency and competitiveness of U.S. small and medium-sized manufacturers

Experience

• Provide engineering students and students in related disciplines with applied experiences not available in the classroom

Pipeline

• Develop a pipeline of energy engineers for the workforce
Overview
Stakeholder Map

- Students
- Faculty
- Manufacturers
- Utilities
- ESCOs

Industrial Assessment Centers
Approximately 50 IAC centers have been supported at different times since 1981.

More than 2,000 students have participated in the program.
Short background
Activities

Types of Clients
- Fabricated Metal Products
- Food and Kindred Products
- Rubber and Miscellaneous Plastics Products
- Industrial and Commercial Machinery and Computer Equipment
- Primary Metal Industries

More than 16,000 IAC clients served

Types of recommendations
- Eliminate leaks in inert gas and compressed air lines/valves
- Utilize higher efficiency lamps and/or ballasts
- Repair and eliminate steam leaks
- Repair or replace steam traps Insulate bare equipment
- Use most efficient type of electric motors
- Use more efficient light sources
- Analyze flue gas for proper air/fuel ratio
- Use multiple speed motors for variable pump, blower and compressor loads
- Preheat boiler makeup water with waste process heat.
Evaluation
Evaluation Design

Workforce Goal of the Impact Assessment

What skills has the program imparted to participants?

What effects on career trajectory?
Evaluation Design
Multi-methods Approach

Impacts on participating students

Context
- Interviews – directors and current students

Efficiency skills
- IAC student alumni survey

Efficiency skills
- Resumes of IAC alumni

Traditional methods

Administrative records method
IAC alumni impacts
Specific, applicable skills in energy efficiency

Types of Energy Efficiency Skills

- Applied experience with a variety of industrial systems and equipment
- Applied experience using energy consumption monitoring and verification equipment
- Experience calculating payback periods related to energy efficiency recommendations
- Experience presenting preliminary energy efficiency recommendations and rationale to maintenance...
- Experience drafting energy efficiency recommendations in reports to clients

Percent of survey respondents

© 2015 SRI International
Resume Analysis

Data Collection

Data Sources

- Direct Contact
  - High quality
  - Unstructured
- Indeed.com
  - High quality
  - Structured
- LinkedIn
  - Low quality
  - Structured
  - Unavailable

Analysis groups

- Participants
  - Random sample
- Cohort
  - same school, degree, graduation date
- Energy Efficiency
  - energy efficiency job holders

Data Collection → Parsing / Ingest → Skill Extraction → Automated Job Classification → Metric Generation
Resume Analysis

Data Collection

Manual Data Collection

- Participant Group
  - Extract contact information from alumni database, contact via e-mail/phone
  - Sample=500 (352 working), 85 + 29 = 109 resumes, 113 surveys

- Cohort Group
  - Query Indeed based on cohort characteristics (received resumes)
  - N=301

Automated Data Collection

- Energy Efficiency Group
  - Build web crawler to extract results of keyword search from Indeed
  - Too many results for one query, break into 50 state-level queries and aggregate results.
  - 5000+
  - Subset for total time in workforce N=867
Resume Analysis

Parsing/Ingest

Develop database schema

No need for relational DB capabilities, so we'll use a document DB instead.

```javascript
skills: [
    "energy auditing",
    "bowling"
],
jobs: [{
    company: "GreenTech, Inc.",
    start: "1/15/2005",
    end: "06/15/2010"},
    {company: "GreenerTech Corp.",
    start: "07/01/2010"
}],
education: . . .
```
title: "Energy Analyst",
description: "Reported directly to Executive Director of Energy Management supporting all departmental energy projects, specifically the 15 MW Riker's Island Cogeneration Project.",
start: 2009-06,
end: 2010-10,
onet: “17-2199.03 - Energy Engineer"
energy_efficiency: True,
Resume Analysis: Skill Extraction

- 9,000 skills and associated salary data set developed by Brookings MPP
- Manually coded 465 of these skills as related to “energy efficiency” activities

Resume skills
- Energy Efficiency
- Energy Audits
- Energy Assessment
- Installation
- Inspection

Data Collection
- Renewable Energy
- Equipment Effectiveness
- Equipment Efficiency

Calculation
- Optimization
IAC alumni impacts
Specific, applicable skills in energy efficiency

Algorithmically tagged resumes with energy efficiency skills and salaries

<table>
<thead>
<tr>
<th></th>
<th>Average Number of Energy Efficiency Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alumni Resumes</td>
<td>8.9</td>
</tr>
<tr>
<td>Energy Resumes</td>
<td>5.5</td>
</tr>
<tr>
<td>Cohort Resumes</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Higher “value” of energy efficiency skills

<table>
<thead>
<tr>
<th></th>
<th>Dollars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alumni Resumes</td>
<td>$72,964</td>
</tr>
<tr>
<td>Energy Resumes</td>
<td>$66,754</td>
</tr>
</tbody>
</table>

Algorithmic processes:
- Data Collection
- Parsing / Ingest
- Skill Extraction
- Automated Job Classification
- Metric Generation
Resume Analysis:
Automated Classification of Career Trajectory

- **ONET Title**
  - 1100 job categories
  - Training / test mismatch 😞
  - Accuracy: 65%

- **Energy efficiency skills**
  - Binary: yes/no
  - Accuracy: 95% 👍
IAC alumni impacts
Career trajectories

Shorter time of entry into first EE Job

Alumni Resumes: 856 days
Energy Resumes: 1643 days

Portion of Total Career Spent in EE

Alumni Resumes: 42%
Energy Resumes: 28%

72% Surveyed alumni said IAC helped them get 1st job
IAC alumni impacts
Expanding the pipeline of energy efficiency engineers

- 60% Joined because interested in energy efficiency
- 40% Joined for a non-energy efficiency related reason

- 70% Said the program made them more interested in energy efficiency
- 40% Went on to have careers in the energy efficiency sector

53% Surveyed alumni’s first job related to energy efficiency
The IAC program had measurable impact on energy efficiency skills developed and alumni career trajectories.

Our resume technique is a rich source for evaluation of human capital investments.
Other SRI ‘Big Data’ projects:

Estimating the size of the energy engineering workforce for the DOE Advanced Manufacturing Office

Identifying the evolution of rare earths research through advanced analytics for the DOE Advanced Manufacturing Office

Study of technology emergence and the evolution of solar energy technologies through advanced analytics for the DOE Sunshot Initiative
Thank You