Big Data Education Issues for Community College Technician Education Programs

Joyce Malyn-Smith, EDC
Lou Piazza, BATEC
Michael Carter, Johnson County Community College
Ryan Murphy, Sinclair Community College
Michael Harris, Bunker Hill Community College

ATE 2015 PI Meeting: Preparing the Technical Workforce through Innovation, Creativity & Practice
Washington, DC
October 22, 2015
Creating Pathways to Big Data Careers

To address the urgent and rapidly growing need for workers proficient in the use of Big Data skills, EDC’s Oceans of Data Institute (ODI) of Education Development Center, Inc. (EDC), Bunker Hill CC, Johnson County CC, Normandale CC and Sinclair CC are partnering to create a career pathway model for Big Data careers.

1. Define duties, tasks, skills, knowledge and abilities needed by middle skilled big data workers

2. Develop and test model pathways to big data careers.
Questions of the day – Setting the Context

• Who works with Big Data and what do they do? – ODIs Profile of the Big Data Enabled Specialist

• How is the emergence of big data impacting local industry partners and community college technician education programs? – BATEC’s Landscape Study – Lou Piazza
Questions of the day: How are community colleges responding to the need for big data workers?

- What strategies are community colleges new using to respond to the emerging importance of big data? And, How does a college that has a smattering of data analytics classes organize a program or certificate on big data? - Mike Carter, Johnson County CC; Ryan Murphy, Sinclair CC

- What do big data career pathways look like? Mike Harris, Bunker Hill CC

- Q&A
Who works with big data and what do they do?

What are the skills, knowledge and behaviors of a “big data-enabled professional”?

- Medical Informatics
- Astrophysics
- Telecommunications
- Utilities
- Law Enforcement/Forensics
- Hydrology
- Climate Modeling
- Education
- Bioacoustics
- Hazard Analysis
- Analytical Journalism
- Marketing
Expert panel- August 14-15, 2014
Work session- August 14-15, 2015
The Big Data Enabled Specialist is an individual who wrangles and analyzes large and/or complex data sets to enable new capabilities including discovery, decision support and improved outcomes.
What do they need to know and be able to do?

What do they do?
• Duties
• Tasks

What enables them to do this successfully?
• Skills
• Knowledge
• Behaviors
• Tools
• Future Trends/ Concerns
Why do we need to know this?

• To design programs leading to big data careers
• To develop curriculum and assessments
• To build articulation agreements
• To design work-based learning/internships
• To engage employers in program activities
• To provide career counseling
• To develop job descriptions/performance reviews
What are the employment opportunities?

Big Data Landscape Study:

Lou Piazza
Director BATEC
Labor Market Analysis: 3 Research Reports

- **Sizing the Middle Skills Employment Gap**
  - Horizontal study of 9 middle skill occupations

- **BIG Data – Analyzing Employment in the Information Economy**
  - Vertical study of the workforce issues driving the information economy

- **Cyber Security – Protecting our Information Assets**
  - Comprehensive analysis of the requisite skills and necessary credentials to start and advance in the hottest and, arguably, most vital employment segment of this decade
Top Eight Findings

1. Fundamental proficiency in Big Data skills is creeping into a broad range of occupations.
   • Domain Expertise in Business, Marketing, Biotechnology, Health Care, Financial Services, Manufacturing

2. Advancement and longevity require a personal commitment to continued education.
   • Data Engineer / Data Analyst / Data Scientist

3. Big Data jobs pay well.
   • Average salary of $80,720 / 25% premium over average for IT.

4. Employers are struggling to fill Big Data jobs.
   • Job Postings remain open 25% to 50% longer than average

5. The Big Data Employment Sector is large and growing.
   • 1.1 million online postings / 6% of all advertised jobs

6. A unique combination of IT and statistical skills are required in this field.
   • Large Volumes of Data / Unstructured Data / Public Datasets

7. The Pathway is built on the back of traditional data management jobs.
   • Data Engineering / Data Management / Information Technology

8. Career Pathway is gaining definition, but there is little employment history.
   • No BLS / O-Net Data available
Description: BIG Data

- **Overview**
  - Vertical study of the trends in Data Science, Data Analysis, and Data Management
  - Comprehensive study of career opportunities and job transitions for 2 Year Graduates, 4 Year Graduates and Post Graduates

- **Contents**
  - Demand (current openings), average salary, 5 year trends
  - Skills, Credentials and Experience (Competencies, Certifications, Education)
  - Typical Work Functions
  - Location Quotients for Certain Regions
  - Pathway and Transition Analysis
## Career Pathway

<table>
<thead>
<tr>
<th>Occupational Category</th>
<th>Description</th>
<th>Job Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scientist</td>
<td>Design and architect advanced analytical models using advanced tools and complex statistics.</td>
<td>Data Scientist, Statistician, Biostatistician, Quantitative Analyst, Chief Analytic Officer</td>
</tr>
<tr>
<td>Data Analyst</td>
<td>Build and manipulate quantitative models that provide analytical insights using a combination of internal data assets and publically available information.</td>
<td>Data Analyst, Research Analyst, Intelligence Specialist</td>
</tr>
<tr>
<td>Data Engineer</td>
<td>Build and maintain the information infrastructure, collecting, storing, maintaining and managing the data elements.</td>
<td>Database Administrator, Information Architect, Data Architect, Data Warehousing Specialist</td>
</tr>
<tr>
<td>Domain Specialist</td>
<td>Use basic tools and pre-built analytical models to gain insight into areas of functional expertise</td>
<td>Business Analyst, Operations Analyst, Market Researcher, Research Associate, Financial Analyst</td>
</tr>
</tbody>
</table>
Pathway Dimensions

- Data Scientists, 22,644, 3%
- Data Analysts, 130,030, 17%
- Engineers, 629,055, 80%
## Key Statistics

<table>
<thead>
<tr>
<th>Role</th>
<th>Total Postings</th>
<th>Growth</th>
<th>Average Salary</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scientist</td>
<td>22,644</td>
<td>94%</td>
<td>$92,044</td>
<td>38 Days</td>
</tr>
<tr>
<td>Data Analyst</td>
<td>130,030</td>
<td>-24%</td>
<td>$76,535</td>
<td>32 Days</td>
</tr>
<tr>
<td>Data Engineer</td>
<td>629,055</td>
<td>23%</td>
<td>$84,601</td>
<td>34 Days</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Role</th>
<th>2 Yr / 4 Yr</th>
<th>Masters</th>
<th>PhD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scientist</td>
<td>21%</td>
<td>25%</td>
<td>55%</td>
</tr>
<tr>
<td>Data Analyst</td>
<td>69%</td>
<td>23%</td>
<td>9%</td>
</tr>
<tr>
<td>Data Engineer</td>
<td>78%</td>
<td>17%</td>
<td>5%</td>
</tr>
</tbody>
</table>
What are the big data issues for community colleges?

- What do people working with big data need to know and do?
- How is the emergence of big data impacting local industry partners and community college technician education programs?
- What strategies are community colleges using to respond to the emerging importance of big data?
- What might big data career pathways look like?
- What challenges are faced by community colleges implementing strategic plans to develop pathways to big data careers?
BHCC Data Management Certificate
In Curriculum Fall 2015

• 16 Credit Certificate
  – Geared at entry level DB associates

• Course Matrix
  – CIT-113 IT Problem Solving or equivalent (3cr)
    • Entry Prerequisite for all CIT courses.
    • Requires developmental Math and English prerequisites
  – CIT-137 Intro to Big Data with R and R Studio (4cr)
  – CIT-162 Intro to Networking (3cr)
  – CIT 234 Decision Support Excel (3cr)
  – CIT-236 SQL Programming (3cr)
BHCC Data Analytics Certificate
In Development Expected FA 16 Launch

• 12 Credit Certificate
  – Data Management Certificate stacks into Data Analytics Certificate
  – 28 Total credits needed (16 from Data Management)

• Course Matrix (all 3-credit courses)
  – CIT-543 Data Analytics and Predictive Analysis (1st semester offered FA15)
  – CIT-125 Python Programming
  – CIT-240 Database Programming or CIT-XXX Hadoop (TBD)
  – MAT-182 Statistics (learning community with CIT-543)
What are the big data issues for community colleges?

- What do people working with big data need to know and do?
- How is the emergence of big data impacting local industry partners and community college technician education programs?
- What strategies are community colleges using to respond to the emerging importance of big data?
- What might big data career pathways look like?
- What challenges are faced by community colleges implementing strategic plans to develop pathways to big data careers?
What are your questions?

Q & A
Joyce Malyn-Smith, PI  jmsmith@edc.org
Mike Carter, Co-PI  mcarter8@jccc.edu
Ryan Murphy, Co-PI  Ryan.Murphy@sinclair.edu
Mike Harris  mdharris@bhcc.mass.edu
Lou Piazza  lou.piazza@umb.edu

For more information