**Developing a Profile of a Big Data-Enabled Specialist: Brief Outline of the Process**

EDC’s approach integrates two methods used in Germany to develop curriculum for technical training programs.: 1) the “training occupation” which defines an occupation based on a shared set of skills and knowledge used by workers in a range of related professions and 2) the DACUM (Developing a Curriculum) occupational analysis model widely used in the US and internationally for more than 50 years. EDC has applied this process to ~ 20-30 projects in different professional contexts to broadly define emerging occupations as well as existing professions undergoing substantive change. Examples of prior projects are:

* Bioscience National Skills Standards
* Information Technology Across Careers
* Computational Thinking in America’s Workplace
* New Media Enabled Technician

EDC’s modified DACUM methodology rests upon three basic principles:

* Expert workers can describe and define their jobs more accurately than anyone else.
* An effective way to define a job is to precisely describe the tasks that expert workers perform.
* All tasks, in order to be performed correctly, demand certain knowledge, skills, resources, and behaviors.

The Profiling Process

1. A panel of 8-10 expert workers is recruited. These experts represent a range of occupational levels and work settings currently involved in data science. They are individuals recognized by their peers as experts and have a minimum of two-years experience in the field.
2. EDC works with the panel on-line to draft an initial occupational definition of a Data Scientist. This is basically a common core of skills used by workers within a range of occupations across an industry sector. This is broader than any specific job an individual might have, and represents all of the shared work tasks, knowledge, skills and attributes required to perform a range of job functions.
3. The panel is convened for a two-day intensive work session to conduct the modified DACUM analysis. During this workshop, EDC facilitates the work of the panel through a structured process as they:
   1. review, discuss, and refine the proposed occupational definition so that it captures the essence and commonalities of their own work in data science.
   2. produce descriptions of concrete, observable work activities performed by data scientists.
   3. organize the activities around major areas of work-related responsibilities and identify: 1) the skills and knowledge necessary to perform these responsibilities; 2) the behavior, or attributes, demonstrated by successful data scientists, and 3)the resources necessary for successful performance of these responsibilities.
   4. identify major industry trends that define the current context for data science.
   5. A draft profile is completed by the end of the 2-day workshop.
4. The profile is validated by a larger community of data scientists using a customized on-line occupational survey.