

# DTL-DISC Course

## Big Data

*tentative date: Fall 2013*

DISC meeting, October 19, 2012  
Irene Nooren  
Celia van Gelder

# DTL-DISC Education of data experts

The new type of DISC data expert should get specific education to become flexible and highly skilled engineers  
*(from DISC twopager)*

## **Step 1:**

Who are the foreseen DISC data experts?

What kind of expertise is needed for Big Data work?

## **Step 2:**

What do they need to learn in a Big Data Course?

## **Step 3:**

Who can train/educate them?

# Big data requires new types of jobs

7 new types of jobs created by Big Data

- Data scientists
- Data architects
- Data visualizers
- Data change agents
- Data engineer / operators
- Data stewards
- Data virtualization/cloud specialists.

<http://www.smartplanet.com/blog/bulletin/7-new-types-of-jobs-created-by-big-data/682>

## **Step 1: Who are the foreseen DISC data experts? What kind of expertise is needed for Big Data work?**

- **Data scientists**
- **Data architects**
- **Data visualizers**
- Data change agents
- **Data engineer / operators**
- Data stewards
- Data virtualization/cloud specialists.
- Other?

## **Step 2 & 3:**

### **What do they need to learn in a Big Data Course?**

### **Who can train/educate them?**

- data architecture en design - ?
- database management - ?
- data analysis and processing (statistics) - WP statistics
- data mining and interpretation - ?
- data integration, context and visualization - ?
- data modeling - WP systems biology
- data sharing and security – SURFsara
- data governance and quality management - ?
- data infrastructure (storage, compute, network) – SURFsara
- data resource management - ? / EBI ea
- other?

# Duration & Course Format

## **Duration & format:**

Has yet to be defined after setting of specific learning goals

**Tentative course date:** Fall 2013

# Who?

- Irene Nooren
- Machiel Jansen
- Celia van Gelder
- You!
- You!
- You!

- The end

# 7 types of jobs (1)

- **Data scientists:** This emerging role is taking the lead in processing raw data and determining what types of analysis would deliver the best results. Typical backgrounds, as cited by Harbert, include math and statistics, as well as artificial intelligence and natural language processing.
- **Data architects:** Organizations managing Big Data need professionals who will be able to build a data model, and plan out a roadmap of how and when various data sources and analytical tools will come online, and how they will all fit together.
- **Data visualizers:** These days, a lot of decision-makers rely on information that is presented to them in a highly visual format — either on dashboards with colorful alerts and “dials,” or in quick-to-understand charts and graphs. Organizations need professionals who can “harness the data and put it in context, in layman’s language, exploring what the data means and how it will impact the company,” says Harbert.

# 7 types of jobs (2)

- **Data change agents:** Every forward-thinking organization needs “change agents” — usually an informal role — who can evangelize and marshal the necessary resources for new innovation and ways of doing business. Harbert predicts that “data change agents” may be more of a formal job title in the years to come, driving “changes in internal operations and processes based on data analytics.” They need to be good communicators, and a [Six Sigma](#) background — meaning they know how to apply statistics to improve quality on a continuous basis — also helps.
- **Data engineer/operators:** These are the people that make the Big Data infrastructure hum on a day-to-day basis. “They develop the architecture that helps analyze and supply data in the way the business needs, and make sure systems are performing smoothly,” says Harbert.

# 7 types of jobs (3)

- **Data stewards:** Not mentioned in Harbert's list, but essential to any analytics-driven organization, is the emerging role of data steward. Every bit and byte of data across the enterprise should be owned by someone — ideally, a line of business. Data stewards ensure that data sources are properly accounted for, and may also maintain a centralized repository as part of a Master Data Management approach, in which there is one “gold copy” of enterprise data to be referenced.
- **Data virtualization/cloud specialists:** Databases themselves are no longer as unique as they use to be. What matters now is the ability to build and maintain a virtualized data service layer that can draw data from any source and make it available across organizations in a consistent, easy-to-access manner. Sometimes, this is called “Database-as-a-Service.” No matter what it's called, organizations need professionals that can also build and support these virtualized layers or clouds.