Who Is Doing Our Data Laundry?  
*
**EDUCAUSE 2017**

**Brad Wheeler, Ph.D.**
IU Vice President for Information Technology, & CIO
Professor of Information Systems, Kelley School of Business
What should I do?
What’s the Campus Imperative?

Rapidly remediate information reporting and dashboards to get everyone on a common basis of facts.

Accelerate Student Success goals in terms of graduation rates, retention, course sequencing, etc.

Benchmark an institution’s performance on key performance indicator ratios for efficiency and effectiveness.

Enable better financial decisions by deans and department chairs as they make choices in degrees to offer, frequency of courses, and teaching assignments.

Empower advisors and students with predictive information to support wise decisions in major and course selection;
Capital $
Consumables $
Who is Doing Our Data Laundry?

The EDUCAUSE Review Online, 13 March 2017

By data laundry, I am referring to the legitimate process of transforming and repurposing abundant data into timely, insightful, and relevant information for another context. It is a mostly unseen, antecedent process that unlocks data’s value and insights for the needs of decision makers.
Why is Our Data Laundry Piling Up?

"Our institutions are often quite data-rich and insight-poor."

Increasing pressures from external and internal forces for better decisions.
Doing the Data Laundry

The data laundry work includes two distinct and sequenced phases. First, data cleaning prepares the data for reuse in a different context....

Clean data likewise needs additional work, and that is done in the second phase: data (information) presentation contextualizes the information into reports, graphs, dashboards, and decision support systems. ”
Data Cleaning

- Discovering
- Extracting
- Re-coding
- Uploading
- Normalizing
Information Presentation

Enriching  Comparing  Presenting
Data Laundry and the Magic Bundle

Learn and Revise Loop

Red = Presumed work and expertise advantage
**Thick** white borders denote challenging tasks
Dashed borders indicate optional steps that may not be in every project
Insource or Outsource?
<table>
<thead>
<tr>
<th>TASKS</th>
<th>IN-HOUSE (OR VIA CONSULTANTS)</th>
<th>OUTSOURCED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovering data sources that support the goal</td>
<td>Knowing where data resides in campus systems</td>
<td>Possibly knowing of external data sources</td>
</tr>
<tr>
<td>Extracting data from systems of record</td>
<td>Conducting considerable functional and technical work to extract data from core systems</td>
<td>Technical consulting</td>
</tr>
<tr>
<td>Re-coding data to preserve context and conform to new formats</td>
<td>Using local expertise or work to convey business logic and to format data</td>
<td>Knowledge of intended uses and models required to adapt context to new purpose</td>
</tr>
<tr>
<td>Uploading</td>
<td>Extensive internal reviews of uses for legal/policy compliance before data is uploaded to campus or vended systems</td>
<td>Technical consulting</td>
</tr>
<tr>
<td>Normalizing</td>
<td>Normalizing data to new context and as inputs for models</td>
<td>Normalized to vendor’s models</td>
</tr>
<tr>
<td>Enriching</td>
<td>Enriching with local predictive models</td>
<td>Enriched with vendor’s predictive models or peer benchmarks</td>
</tr>
<tr>
<td>Comparing</td>
<td>N/A or internal within/across an institution</td>
<td>Enriched with normalized peer group comparisons</td>
</tr>
<tr>
<td>Presenting</td>
<td>Providing dashboards, screens, reports</td>
<td>Dashboards, screens, reports</td>
</tr>
<tr>
<td>Workflow (often never achieved)</td>
<td>Integrating the new presentations into workflow and other campus systems</td>
<td>Technical consulting</td>
</tr>
</tbody>
</table>
While working at Tesla, I always enjoyed talking to people after they finished a factory tour. As much as they raved about the amazing automation, gigantic presses, and hundreds of robots, the reality was they only saw half of the actual manufacturing that was taking place in the building.
IU’s Data Laundromat

2015...
Indiana University is a data-rich organization, as it collects elements on all aspects of students, faculty and staff, facilities, finances, budgets and research. However, this assessment has uncovered that although all of this information is collected, it is not in the most usable format as possible for the key decision makers at all levels within the university…. the university is lacking in its ability to perform analysis across the functional areas both in a structured format as well as in an ad-hoc nature.
IU Decision Support Initiative (2015)

1. Enable report and dashboard “discovery” via search
   ds.iu.edu  Decision Support

2. Created a factory for Decision Support Initiatives

3. Agile methodology -- then run, run, run!
Decision Support Initiative Goal

“To improve decision making at all levels of Indiana University by dramatically enhancing the availability of timely, relevant, and accurate information to support decision makers”

http://decisionsupport.iu.edu
DSI will affect People and Organizations. It addresses changes in culture, roles, authority over data, access, skills, work habits...

DSI will affect Data and Technology. Over time, it will address standardizing Data Definitions and change some of the core Data, Information, and Reporting IT.

DSI will affect departments, schools, and campuses across IU. It will provide greater transparency into and across many areas. The focus is to improve decision outcomes for IU in a rapidly changing world.
Insourced Discovery and Dashboards

ds.iu.edu
Insourced Tools + Outsourced Services

Via the Unizin Consortium
DSI Project Pipeline

Many ideas…

Only formal Project Charters advance for consideration and prioritization.

Charters detail merits of the project, scope, commitments of functional staffing, etc.

Charters Queue:
- Project Charter
- Project Charter
- Project Charter
...

DSI Steering Council
- Assesses resource capacity
- Approves Charters
- Prioritizes Charters

Work Queue:
1. Approved Charter
2. Approved Charter
3. Approved Charter
...

DSS Factory
Disciplined, Formal Agile Process (explained in slide #x)

Agile Team #1
Agile Team #2
Agile Team #3

Functional and Technical Staff, Highly Prescriptive Process, Time-boxed 2 Week Delivery Sprints

http://ds.iu.edu
- Reports
- Dashboards
- Decision Support

• Education
• Training
• Rollout
• Diffusion
• Evaluation
• Support
• Enhancement

+
IU Academic Metrics 360

- Credit Hours
- HR data
- Financials
- Trends and forecasts
- Tailored to *How we make decisions*
- Data Laundry Work!!
Will clean data lead to good decisions?
1. **Give them more information and decision making will improve.**
   
   *In many cases, decision makers suffer from an overload of irrelevant information more than a lack of relevant information for a decision.*

2. **Decision makers know what information they need.**
   
   *In many cases, decision makers often request vastly more information than needed to optimally make a decision (exacerbating #1).*

3. **Give decision makers the information that they need and decisions will improve.**
   
   *In many cases, decision makers do not have refined models of the relevant information that best predict desired outcomes.*

---

**Caution!**

What’s up with YOUR data laundry?
Who Is Doing Our Data Laundry?

*EDUCAUSE 2017*

**Brad Wheeler, Ph.D.**

IU Vice President for Information Technology, & CIO
Professor of Information Systems, Kelley School of Business

October 2017
Data Laundry and the Magic Bundle

Learn and Revise Loop

Red = Presumed work and expertise advantage
**Thick** white borders denote challenging tasks
Dashed borders indicate optional steps that may not be in every project