Moving From “Drowning in the Demand for Data” to Smooth Sailing

Barbara A. Knuth
Dean of the Graduate School
Cornell University

CGS Session: Navigating the Sea of Surveys, December 2018
“Drowning in the Demand for Data”

• Growing expectations regarding data collection & transparency
  • External
  • Internal

• Strategizing and prioritizing with limited resources
  • Identify potential goals, uses, impacts
  • Identify potential audiences

• Achieving impact through data
  • Improve programs
  • Influence decisions (e.g., resource allocation, resource capture)

• Local & national context and implications
Using Data Effectively & With Impact @ Cornell:

• Annual report to the provost
  • Document change over time
  • Progress made
  • Areas for improvement, resource needs

• Biennial internal program review meetings

• Interactive dashboards for faculty directors of graduate students & staff assistants

• Public interactive dashboards & reports:
  • Informed prospective students
  • Transparency for enrolled students
  • Accountability by graduate programs

• Situating ourselves in national conversations
Cornell Example: Doctoral Career Outcomes vs. National Narratives (shift away from TT, into non-academic)

- 2-20 years out:
  - Government consistently 6-8%
  - Business consistently 31-34%
  - Education Non-TT 10-14%
  - Education TT 41-48%
- Outliers:
  - 12 yrs out (??)
  - 2 yrs out (early career formation? Or sea change?)
Cornell Example: Doctoral Career Outcomes

### Employment Sectors by Years Post-Degree (6,040)

**All Fields**

<table>
<thead>
<tr>
<th>Years</th>
<th>Self-employed, Other</th>
<th>Government</th>
<th>Business, Industry, Non-Profit</th>
<th>Education Non-Tenure-Track</th>
<th>Education Tenure-Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 yrs</td>
<td>5%</td>
<td>7%</td>
<td>33%</td>
<td>45%</td>
<td>10%</td>
</tr>
<tr>
<td>18 yrs</td>
<td>8%</td>
<td>8%</td>
<td>31%</td>
<td>46%</td>
<td>10%</td>
</tr>
<tr>
<td>15 yrs</td>
<td>8%</td>
<td>8%</td>
<td>34%</td>
<td>46%</td>
<td>12%</td>
</tr>
<tr>
<td>12 yrs</td>
<td>6%</td>
<td>8%</td>
<td>33%</td>
<td>46%</td>
<td>10%</td>
</tr>
<tr>
<td>10 yrs</td>
<td>6%</td>
<td>8%</td>
<td>33%</td>
<td>46%</td>
<td>10%</td>
</tr>
<tr>
<td>7 yrs</td>
<td>6%</td>
<td>8%</td>
<td>33%</td>
<td>44%</td>
<td>10%</td>
</tr>
<tr>
<td>5 yrs</td>
<td>8%</td>
<td>8%</td>
<td>34%</td>
<td>41%</td>
<td>10%</td>
</tr>
<tr>
<td>2 yrs</td>
<td>6%</td>
<td>8%</td>
<td>34%</td>
<td>41%</td>
<td>10%</td>
</tr>
</tbody>
</table>

### Humanities & Arts

<table>
<thead>
<tr>
<th>Years</th>
<th>Self-employed, Other</th>
<th>Government</th>
<th>Business, Industry, Non-Profit</th>
<th>Education Non-Tenure-Track</th>
<th>Education Tenure-Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 yrs</td>
<td>7%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>11%</td>
</tr>
<tr>
<td>18 yrs</td>
<td>14%</td>
<td>14%</td>
<td>11%</td>
<td>13%</td>
<td>12%</td>
</tr>
<tr>
<td>15 yrs</td>
<td>13%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>12 yrs</td>
<td>14%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>10 yrs</td>
<td>8%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>7 yrs</td>
<td>7%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>5 yrs</td>
<td>0%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>2 yrs</td>
<td>9%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
</tbody>
</table>

### Life Sciences

<table>
<thead>
<tr>
<th>Years</th>
<th>Self-employed, Other</th>
<th>Government</th>
<th>Business, Industry, Non-Profit</th>
<th>Education Non-Tenure-Track</th>
<th>Education Tenure-Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 yrs</td>
<td>18%</td>
<td>14%</td>
<td>10%</td>
<td>10%</td>
<td>11%</td>
</tr>
<tr>
<td>18 yrs</td>
<td>14%</td>
<td>10%</td>
<td>11%</td>
<td>13%</td>
<td>12%</td>
</tr>
<tr>
<td>15 yrs</td>
<td>10%</td>
<td>10%</td>
<td>11%</td>
<td>13%</td>
<td>12%</td>
</tr>
<tr>
<td>12 yrs</td>
<td>10%</td>
<td>10%</td>
<td>11%</td>
<td>13%</td>
<td>12%</td>
</tr>
<tr>
<td>10 yrs</td>
<td>10%</td>
<td>10%</td>
<td>11%</td>
<td>13%</td>
<td>12%</td>
</tr>
<tr>
<td>7 yrs</td>
<td>11%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>5 yrs</td>
<td>27%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>2 yrs</td>
<td>9%</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>10%</td>
</tr>
</tbody>
</table>

### Physical Sciences

<table>
<thead>
<tr>
<th>Years</th>
<th>Self-employed, Other</th>
<th>Government</th>
<th>Business, Industry, Non-Profit</th>
<th>Education Non-Tenure-Track</th>
<th>Education Tenure-Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 yrs</td>
<td>5%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>18 yrs</td>
<td>7%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>15 yrs</td>
<td>4%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>12 yrs</td>
<td>4%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>10 yrs</td>
<td>8%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>7 yrs</td>
<td>8%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>5 yrs</td>
<td>8%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>2 yrs</td>
<td>6%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
</tbody>
</table>

### Social Sciences

<table>
<thead>
<tr>
<th>Years</th>
<th>Self-employed, Other</th>
<th>Government</th>
<th>Business, Industry, Non-Profit</th>
<th>Education Non-Tenure-Track</th>
<th>Education Tenure-Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 yrs</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>18 yrs</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>15 yrs</td>
<td>7%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>12 yrs</td>
<td>7%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>10 yrs</td>
<td>6%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>7 yrs</td>
<td>6%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>5 yrs</td>
<td>5%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>2 yrs</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
</tr>
</tbody>
</table>
Cornell’s Multi-institutional Data-Related Engagements

Program Evaluation and Improvement
• AGEP (diversity for the future professoriate, campus climate)
• BEST (career exploration)
• CIRTL (preparing future faculty)

“Private” Collaborative Data Exchange
• AAUDE (data comparisons)

“Public” Data Transparency
• CNGLS (graduate students & postdocs)
• Council of Graduate Schools (enrollment, international, etc.)
• SED, etc.

Median Time-to-Degree by Discipline Comparison

- Humanities
- Life Sciences
- Physical Sciences
- Social Sciences

- Cornell
- AAUDE Doctoral Research Schools
- AAUDE Ivy Plus Schools
- SED national data
Multi-institutional Benefits

+ Generate comparative data (vs. peers)
+ Contextualize graduate education (nationally)
+ Share what works for evidence-based improvement
+ Reduce risk from unilateral data transparency
Multi-institutional Benefits and Challenges

+ Generate comparative data (vs. peers)
+ Contextualize graduate education (nationally)
+ Share what works for evidence-based improvement
+ Reduce risk from unilateral data transparency

– Agree to shared or flexible data definitions
– Reconfigure data for multiple internal and external needs
– Factor in existing data collected prior to multi-institution agreements
– Reconcile lumpers vs. splitters
– Coordinate data across different institutional structures
  • internally (e.g., Graduate Schools vs. Postdoc Offices vs. HR vs. IR)
  • externally (e.g., CIP code variation, data sharing restrictions vs. sunshine laws)
Weighing Pros & Cons of Participation

• What is the potential benefit?
  • To my graduate school
  • To my institution
  • To my students/faculty/staff
  • To higher education

• How much will it cost to participate?
  • $$
  • Staff time
  • Student/faculty time

• What are the costs of not participating?
  • Institutional reputation, visibility
  • Informed decisions: individually, locally, regionally, nationally
“Private” Data Example:

AAUDE: The Association of American Universities Data Exchange
- AAU institutions
- Participate in exchanging data/information to support decision-making for graduate education at institution
- Graduate education data are not public
- If you submit data, you have access to data

✓ Time to degree
✓ Degree completion rate
✓ PhD exit survey
✓ PhD career outcomes

- Well-developed protocols and definitions for each
AAUDE Experience:

✓ Data protocols explicitly consider compatibility issues:
  • Clear definition for each data element
  • Well-documented data dictionary
  • Compare and contrast similar data elements commonly available
  • Discipline crosswalks
  • Relation to IPEDS, SED, CGS and other data collection systems

✓ Rules for small cell sizes; safeguard individual privacy

✓ Rules to enable meaningful aggregation

✓ Not for rankings

✓ Peer comparisons
  • Internally: group member names only, not individual schools
  • Externally: masked, group name only
Multi-institution Public Data Example: Coalition for Next Generation Life Science

http://nglscoalition.org/
Multi-Institution Data Transparency Example: Coalition for Next Generation Life Science (CNGLS)

• Voluntary agreement by universities and research institutes (26+) to address calls for increased transparency re: life science trainees
  • Graduate students
  • Postdoctoral scholars

• Post data using common definitions (by demographic groups)
  • Admissions
  • Matriculation
  • MTTD and MT in postdoc status
  • Completion
  • Career outcomes (taxonomy by job sector & career type)
Multi-Institution Data Transparency Example:
Coalition for Next Generation Life Science (CNGLS)

- Voluntary agreement by universities and research institutes (26+) to address calls for increased transparency re: life science trainees
  - Graduate students
  - Postdoctoral scholars

- Post data using common definitions (by demographic groups)
  - Admissions
  - Matriculation
  - MTTD and MT in postdoc status
  - Completion
  - Career outcomes (taxonomy by job sector & career type)

- **BUT:**
  - Most graduate schools have responsibilities for multiple disciplines
  - Many graduate schools don’t curate postdoctoral scholar data
  - Some graduate schools already have internal & other partnership protocols for these data
Cornell Example: CNGLS Data

- Posted on websites:
  - Graduate School Academics Info
  - Graduate School Catalog of Degree Programs
  - Individual Degree Programs
  - Office of Postdoctoral Studies (postdoc data)

- No required format
- No single posting location to compare schools

https://tableau.cornell.edu/views/CornellUniversityGraduateSchoolDoctoralProgramStatistics/TheOneDashboard?iframeSizedToWindow=true&:embed=y&:showAppBanner=false&:display_count=no&:showVizHome=no
Cornell Example: CNGLS – graduate student demographics

Filter by Graduate Field or Discipline

(CIP codes may differ among institutions)
How selective is the program?

- Admit Rate
- Yield
- Applications
- Admittances
- Matriculations

<table>
<thead>
<tr>
<th>Year</th>
<th>Admit Rate</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>22%</td>
<td>45%</td>
</tr>
<tr>
<td>2014</td>
<td>14%</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>17%</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>23%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Applications</th>
<th>Admittances</th>
<th>Matriculations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>291</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>2014</td>
<td>277</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2015</td>
<td>291</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2016</td>
<td>279</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2017</td>
<td>298</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
Cornell Example: CNGLS graduate students

Who is in the program?
- Gender
- Citizenship
- Ethnicity
Cornell Example: CNGLS graduate students

How many graduate from the program?

- Completed
- In Progress
- Withdrawn
- Gender
- Citizenship
- Ethnicity
How long does the program take?

- Median Time-to-Degree
- Filter by Graduating Classes
- Filter by Demographics
Flexible Data Visualization Approaches in Coalition: UCSF
Flexible Data Visualization Approaches in Coalition:

UCSF

MIT
Flexible Data Visualization Approaches in Coalition:

UCSF

MIT

PENN
Flexible Data Visualization Approaches in Coalition:

UCSF

MIT

PENN

UMBC
Cornell Example: CNGLS Postdoc Demographics

- CNGLS caused us to look at, and collect, data we hadn’t before
- Required collaboration with other campus offices

- Postdocs:
  - Sex
  - Race/ethnicity
  - Citizenship
  - Time in position
  - Median time to departure

- Coming Soon: Job after departure
Lessons Learned

• Opportunity for comparative data to inform decisions & resource requests

• Promote internal communications at multiple institutional levels

• Share what works and build on progress

• Collaboration may result in better outcomes
  • Data visualizations
  • Provisions for aggregation vs. specificity
  • Data masking as appropriate
  • Consensus, within reason, on data definitions

• Contextualize graduate education in national conversations

• Flexibility and adaptability are important
  • Multiple data transparency efforts for different purposes
  • Real costs of managing multiple efforts