TEN THINGS I HAVE LEARNED
About Health Data Monetization

Khaled El Emam
1. DATA SUPPLY

cc: Nick.Warner - https://www.flickr.com/photos/39708982@N04
Dataset Search

Search for Datasets

Try boston education data or weather site:noaa.gov
REAL WORLD DATA

cc: keith ellwood - https://www.flickr.com/photos/76377775@N05
2. VALUE OF DATA
INCREASING VALUE

• Complement your data from other commercial or public sources
• Commercialize information products not raw data
• Improve data quality
3. VALUE TO PATIENTS

cc: Proctor Archives - https://www.flickr.com/photos/45469741@N05
Data protection

Revealed: how drugs giants can access your health records

Experts say information sold on by Department of Health and Social Care can be traced back to individual medical records
Medical-record software companies are selling your health data
RETURNS TO PATIENTS
4. SOCIAL LICENSE
1. Transparency

2. Data for good

3. Oversight on data uses

4. Pro-active engagement with regulators
5. DATA QUALITY
6. DATA MARKETS
7. Privacy Enhancing Technologies (PETs)
HIPAA Privacy Rule De-identification Methods

Expert Determination § 164.514(b)(1)
- Apply statistical or scientific principles
- Very small risk that anticipated recipient could identify individual

Safe Harbor § 164.514(b)(2)
- Removal of 18 types of identifiers
- No actual knowledge residual information can identify individual
Practical Synthetic Data Generation
Balancing Privacy and the Broad Availability of Data

Khaled El Emam, Lucy Mosquera & Richard Hopcroft
Governance
8. DATA OWNERSHIP
9. Nightmare Scenarios
10. ORGANIZATIONAL DESIGN
% of respondents

Organizations’ response to changes brought about by data and analytics

<table>
<thead>
<tr>
<th>Response Description</th>
<th>Analytics leaders, n = 85</th>
<th>Analytics laggards, n = 83</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altered longer-term corporate strategy</td>
<td>32</td>
<td>18</td>
</tr>
<tr>
<td>Developed coordinated plan but have not altered longer-term corporate strategy</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>Responded through ad hoc initiatives and actions</td>
<td>48</td>
<td>51</td>
</tr>
<tr>
<td>Not yet responded</td>
<td>2</td>
<td>23</td>
</tr>
</tbody>
</table>

Organizational structure of data-and-analytics activities

<table>
<thead>
<tr>
<th>Structure Description</th>
<th>Analytics leaders, n = 85</th>
<th>Analytics laggards, n = 83</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid (led by center of excellence)</td>
<td>45</td>
<td>13</td>
</tr>
<tr>
<td>Hybrid (led by business units)</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Decentralized (analytics employees embedded in business units)</td>
<td>24</td>
<td>21</td>
</tr>
<tr>
<td>Centralized (single corporate center of excellence for analytics activities)</td>
<td>16</td>
<td>24</td>
</tr>
</tbody>
</table>

% of executive teams’ time spent discussing data-and-analytics activities

<table>
<thead>
<tr>
<th>Time Spent Discussing</th>
<th>Analytics leaders, n = 85</th>
<th>Analytics laggards, n = 83</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;50%</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>31%-50%</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>11%-30%</td>
<td>49</td>
<td>32</td>
</tr>
<tr>
<td>≤10%</td>
<td>10</td>
<td>65</td>
</tr>
</tbody>
</table>