Data analytics as inputs for innovations in engineering design

Ma Lorena Escandón Quintanilla
PhD Candidate, ETS & HEC Montreal
Agenda

• Hypothesis
• KDD
• EDP
• Summary of cases
• Findings
Hypothesis

1. Data mining can be used at different stages of the creative process

2. Use of data as input for creativity will increase diversity of ideas

3. Access to data exploration tool will positively impact the generation of “novel combinations”
Knowledge discovery from databases

Piatetsky-Shapiro, 1991

Lorena Escandon, 2017
Engineering design process

1. Problem scoping:
   - Need identification
   - Problem definition
   - Information gathering

2. Develop alternative solutions:
   - Evaluation
   - Feasibility analysis
   - Modeling
   - Idea generation

3. Project realization:
   - Selection
   - Communication
   - Implementation

Atman et al., 2007
### Summary of cases

<table>
<thead>
<tr>
<th>Case</th>
<th>Objective</th>
<th>Duration</th>
<th>Number of teams</th>
<th>Total participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Need identification</td>
<td>4 hours</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Benchmark Novel concept Prototype</td>
<td>≈ 100 hours</td>
<td>8</td>
<td>49</td>
</tr>
<tr>
<td>3</td>
<td>Novel concepts</td>
<td>24 hours</td>
<td>34</td>
<td>≈212</td>
</tr>
</tbody>
</table>
Context

Innovation contests

- Media
- Organizer
- Replication
- Community functionality
- Contest period
- Sponsors / partners
- Marketing / Activation
- Contest phases
- Participation as
- Target group
- Evaluation
- Degree of elaboration
- Reward / motivation
- Facilitation
- Task specificity

Lorena Escandon, 2017
Case 1

Need identification / Problem definition
Case 1
Need identification / Problem definition

Teams:
• identified elements of the problem world
• generated ideas
• were provided access to data visualization
• generated additional ideas (in less time)
Case 1
Need identification / Problem definition

- Engineering design process

- Knowledge discovery from databases

Lorena Escandon, 2017
## Case 1
**Need identification / Problem definition**

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of participants</th>
<th>Issues identified originally</th>
<th>Issues identified with support</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>
Case 2
Development of alternative solutions
Case 2
Development of alternative solutions

• Teams
  – were trained in the use of the tool
  – were given access to tool with pre-loaded data
  – use of tool was optional

• Teams with access to tool
developed ideas further and
were better graded
Case 2
Development of alternative solutions

- Engineering design process

<table>
<thead>
<tr>
<th>Actor</th>
<th>Need identification</th>
<th>Problem definition</th>
<th>Information gathering</th>
<th>Idea generation</th>
<th>Modeling</th>
<th>Feasibility analysis</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>External</td>
<td>□</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Researcher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Knowledge discovery from databases

<table>
<thead>
<tr>
<th>Actor</th>
<th>Data selection</th>
<th>Data pre-processing</th>
<th>Data transformation</th>
<th>Data mining</th>
<th>Interpretation and evaluation</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Case 2
Development of alternative solutions

<table>
<thead>
<tr>
<th>Team</th>
<th>Initial concept</th>
<th>Final concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cone-shaped turbine which gradually removes garbage of different sizes and stores them in a removable container.</td>
<td>Cone-shaped turbine which gradually removes garbage of different sizes and stores them in a removable container, paired with container changing station in the shore.</td>
</tr>
<tr>
<td>2</td>
<td>Floating waste collector set that traps garbage while water continues to flow through (attached to buoys to float).</td>
<td>Floating waste collector cone that traps garbage while water continues to flow through (attached to buoys to float). A turbine in the pipes keeps the flow of water and waste.</td>
</tr>
<tr>
<td>3</td>
<td>Solar-powered set of interconnected buoys which use a Peristaltic pump to send off-shore samples to the shore for collection in QR coded containers.</td>
<td>Solar-powered set of interconnected buoys which use a Peristaltic pump to send off-shore samples to the shore for collection in QR coded containers.</td>
</tr>
<tr>
<td>4</td>
<td>Autonomous ball-like device that measures in situ parameters and collect waste samples.</td>
<td>Ball-like submarine device with barcoded expandable sample container and portable testing unit.</td>
</tr>
<tr>
<td>5</td>
<td>System of wave-breaking pillars that bend under boats and generate electricity from the movement with piezoelectric generators.</td>
<td>System of wave-breaking pillars that bend under boats and generate electricity from the movement with piezoelectric generators.</td>
</tr>
</tbody>
</table>

Lorena Escando
Case 3

Idea generation
Case 3

Idea generation

• Identified “twin challenges”
• Provided keywords from the data analysis from the beginning
• Teams with keywords generated more diverse solutions
Case 3
Idea generation

- Engineering design process

- Knowledge discovery from databases
Case 3

Idea generation
Case 3
Idea generation

Type A Challenges

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Challenge 1</th>
<th>Challenge 2</th>
<th>Challenge 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>6</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Energy</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Type B Challenges

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Challenge 2</th>
<th>Challenge 3</th>
<th>Challenge 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Data sources</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Medium</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>User engagement</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
Findings

1. In all cases, teams using data as input performed better compared to teams not using it

2. Teams using data showed wider diversity in proposed ideas

3. Teams with no support resort to known solutions

4. Direct access to the tool actually decreased the benefits of data access
Conclusions

- It appears to be better to artificially limit the exploration space
- There is value in the combination of existing knowledge + KDD
- Need to align use of data with stage & objective of EDP
THANK YOU!

ma-lorena.escandon-quintanilla.1@etsmtl.net
# Engineering design process

<table>
<thead>
<tr>
<th>Case</th>
<th>Actor</th>
<th>Need identification</th>
<th>Problem definition</th>
<th>Information gathering</th>
<th>Idea generation</th>
<th>Modeling</th>
<th>Feasibility analysis</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>External Researcher Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Internal Researcher Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Internal Researcher Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lorena Escandon. 2017
## Knowledge discovery from databases

<table>
<thead>
<tr>
<th>Case</th>
<th>Actor</th>
<th>Data selection</th>
<th>pre-processing</th>
<th>Data transformation</th>
<th>Data mining</th>
<th>Interpretation and evaluation</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>Researcher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 2</td>
<td>Researcher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 3</td>
<td>Researcher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>