ANES 2008 Election Survey

Testing an Automated Coding Model

Dean Kotchka
Partner
Language Logic, LLC
Language Logic ???

- Founded in 1999
- Publish Verbatim Management software
- Used by over half Honomichl 50 firms
- 12,000 users in 56 countries
- Will process 300 million comments this year
- Includes a ‘Solutions’ Group
Solutions Group

• Provides coding, translating services
• 100+ FTE contractors
• Will process 20 million comments this year
• Used for ANES coding project
The ‘open end’ Dilemma

Balance factors to gain meaningful insight
The ‘open end’ Dilemma

- Cost
- Speed
The ‘open end’ Dilemma

- Cost
- Speed
- Accuracy
- Consistency
One Solution

- Automated Coding Model (ACM)
- Learn by Example approach
- Based on work of Dr. Fabrizio Sebastiani
  - Institute of Information Science and Technologies (ISTI)
- Published papers available
ACM and ANES 2008 Data

• Selected Two Questions

• ‘Dislikes’ – Republican and Democrat

• Question Text:
  Is there anything in particular that you DISLIKE about the REPUBLICAN PARTY?
The ANES Codeframe

1: General
2: Other party
3: Specific Person
4: Ideology/Philosophy
5: Electability
6: Scandal/Cover-up
7: Campaign
8: Other past activity
9: Ability
10: Honesty
11: Intelligence
12: Leadership
13: Management
14: Religion
15: Education
16: Physical Appearance
17: Demographics
18: Policy-Economic
19: Policy-Poor people
20: Policy-Liberty
21: Policy-Enemy countries
22: Policy-Friendly countries
23: Policy – General foreign policy
24: Policy-Other
25: Groups
26: Emotions/Feelings
27: Party-Other
28: Non-Party
29: Party membership
30: Don’t know
31: Refuse
32: Respondent-Other
33: All other
The ANES Codeframe

1: General
2: Other party
3: Specific Person
4: Ideology/Philosophy
5: Electability
6: Scandal/Cover-up
7: Campaign
8: Other past activity
9: Ability
10: Honesty
11: Intelligence
12: Leadership
13: Management
14: Religion
15: Education
16: Physical Appearance
17: Demographics
18: Policy-Economic
19: Policy-Poor people
20: Policy-Liberty
21: Policy-Enemy countries
22: Policy-Friendly countries
23: Policy - General foreign policy
24: Policy-Other
25: Groups
26: Emotions/Feelings
27: Party-Other
28: Non-Party
29: Party membership
30: Don’t know
31: Refuse
32: Respondent-Other
33: All other

ascibe words to wisdom
Republican Dislikes

• ‘Reconciled’ coded comments 2,914
• Selected for ‘training examples’ 1,489
• Comments ‘coded’ through model 1,425
Republican Dislikes

- Comments ‘run’ through model: 1,425
- Comments coded: 1,137
- Comments coded identically: 737
- Really need to view at the code level
Republican Dislikes

• Reconciled codes applied (base) 1,734
• ACM codes applied 1,618
• ACM codes applied not in base 271
  • Precision = 86.4%
• ACM codes missed 368
  • Recall = 82.4%
• Overall Accuracy (F1) 84.3%
<table>
<thead>
<tr>
<th>Comment</th>
<th>ACM (red=ACM added)</th>
<th>Reconciled (blue = ACM Missed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>all political parties are out of touch with American families</td>
<td>General</td>
<td>General</td>
</tr>
<tr>
<td></td>
<td>Policy-Other</td>
<td>Other party</td>
</tr>
<tr>
<td></td>
<td>Party-Other</td>
<td>Groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Party-Other</td>
</tr>
<tr>
<td>Respondent feels the Democratic party is guilty of that being out of</td>
<td>General</td>
<td>General</td>
</tr>
<tr>
<td>touch with American families too</td>
<td></td>
<td>Groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Party-Other</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Respondent-Other</td>
</tr>
<tr>
<td>allowing NAFTA to be voted in in the first place</td>
<td>Policy-Economic</td>
<td>Policy-Economic</td>
</tr>
<tr>
<td></td>
<td>Policy-Liberty</td>
<td></td>
</tr>
<tr>
<td>sometimes they could put forth a little more effort with things going</td>
<td>Groups</td>
<td>Ability</td>
</tr>
<tr>
<td>on with the poor people</td>
<td>Party-Other</td>
<td>Policy-Poor people</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Groups</td>
</tr>
<tr>
<td>I would like them to be more thorough about the programs that they</td>
<td>Policy-Other</td>
<td>Ability</td>
</tr>
<tr>
<td>initiate</td>
<td>Respondent-Other</td>
<td>Policy-Other</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Respondent-Other</td>
</tr>
<tr>
<td>I would like them to make sure that they do what they are supposed to</td>
<td>Respondent-Other</td>
<td>Ability</td>
</tr>
<tr>
<td>do</td>
<td></td>
<td>Respondent-Other</td>
</tr>
<tr>
<td>they are pro homosexuality</td>
<td>Policy-Liberty</td>
<td>Groups</td>
</tr>
<tr>
<td></td>
<td>Groups</td>
<td></td>
</tr>
<tr>
<td>they are pro same sex marriages</td>
<td>Policy-Liberty</td>
<td>Policy-Liberty</td>
</tr>
<tr>
<td></td>
<td>Groups</td>
<td></td>
</tr>
<tr>
<td>too giving</td>
<td>Policy-Economic</td>
<td>Party-Other</td>
</tr>
<tr>
<td>promises that they make</td>
<td>Party-Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Honesty</td>
<td>Party-Other</td>
</tr>
<tr>
<td>I think there should not be a contraceptive issue</td>
<td>Policy-Other</td>
<td>Policy-Liberty</td>
</tr>
<tr>
<td>They could do more</td>
<td>Respondent-Other</td>
<td>Policy-Liberty</td>
</tr>
<tr>
<td>They don't have cohesiveness</td>
<td>Party-Other</td>
<td>Ability</td>
</tr>
<tr>
<td></td>
<td>Ability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leadership</td>
<td>Leadership</td>
</tr>
</tbody>
</table>
ACM / ANES Opportunities

• Formalize Experiment protocol
• Implement full iterative process
• Apply across questions
Thank you

Dean Kotchka
Partner
dean.kotchka@goascribe.com