Test-Driven Development for Technology Policy

Neal Parikh, Ginny Fahs, Brandie Nonnecke
Aspen Tech Policy Hub
The Aspen Institute

November 2019
Technology Policymaking

• Goal: Help make technology policy more robust by
  • Ensuring policy actually addresses problem of interest
  • Avoiding unintended consequences
    • Reducing ‘whack-a-mole’ problem where subsequent drafts of policy introduce conflicts with previously agreed upon goal
  • Adapt computer science and engineering methods used to build and deploy robust, reliable software to policymaking
## Software Tests

<table>
<thead>
<tr>
<th>What’s Being Tested</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal US number, format 1</td>
<td>(650) 723-2300 is accepted</td>
</tr>
<tr>
<td>Normal US number, format 2</td>
<td>650-723-2300 is accepted</td>
</tr>
<tr>
<td>US number in invalid format</td>
<td>65-07-232-300 is rejected</td>
</tr>
<tr>
<td>UK number</td>
<td>+44-20-7925-0918 is rejected</td>
</tr>
<tr>
<td>Special number</td>
<td>911 is rejected</td>
</tr>
<tr>
<td>Non-numeric input</td>
<td>“ABCDEF” is rejected</td>
</tr>
</tbody>
</table>
1. WRITE TEST

2. REWRITE CODE

3. RUN TESTS

Test Passes?

More Cases?

SHIP CODE
Testing for Policy

- “Test” = snippet of text that describes concrete situation the policy is or is not supposed to address

- “Code” = policy language being developed (legal or less formal language)

- “Running the tests” = policy staff looking through a test suite and comparing to current policy language

- Write tests first (feels counterintuitive at first)
Example

- Goal of policy: ban use of “dark patterns” (manipulative online marketing) by “large online platforms” with over 100 million monthly users

- Deceptive Experiences to Online Users Reduction (DETOUR) Act introduced in 2019

Source: https://twitter.com/MarkWarner/status/1115660831969153025
Testing DETOUR Act

Potential Test: don’t allow hiding opt-in / consent language in dropdown box

Source: https://twitter.com/MarkWarner/status/1115660831969153025
Testing DETOUR Act

Potential Test: don’t allow fake hair to make people tap an ad on mobile device

Source: https://twitter.com/MarkWarner/status/1115660831969153025
Testing DETOUR Act

**Potential Test**: allow normal online banner ad for NYT Cooking
Designing Tests

- Policy testing requires creativity, as does software testing
- But there are patterns in types of tests that are useful
- Tests can be reused across policies and the policymaker becomes faster with experience
<table>
<thead>
<tr>
<th>Test Category</th>
<th>Questions</th>
<th>Test Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeric Thresholds</td>
<td>- How does the policy change as numbers in the policy are adjusted?</td>
<td>Covering online platforms with over 1M vs 100M vs 1B active users</td>
</tr>
<tr>
<td></td>
<td>- What happens to entities that pass in/out of the threshold?</td>
<td></td>
</tr>
<tr>
<td>Subcategory</td>
<td>What are all the distinct subcategories of a category referenced in the policy?</td>
<td>For content moderation policy, nudity could include pornography, historical photos, photojournalism, nudes in art</td>
</tr>
<tr>
<td>Valid Practices</td>
<td>If some behavior is being banned, what is some similar but allowed behavior?</td>
<td>Consider valid examples of online marketing or benign A/B testing in addition to dark patterns</td>
</tr>
<tr>
<td>Demographics</td>
<td>Does the policy make sense as you vary the attributes of the people/companies covered?</td>
<td>- Demographics (age, gender, income, location), access to internet, -Market cap, funding stage, sector</td>
</tr>
<tr>
<td>Jargon</td>
<td>How do different people (engineers, domain experts, laypeople) interpret special jargon in the policy?</td>
<td>DETOUR Act includes term “behavioral experiment”, which could be interpreted differently by engineers, users, social scientists</td>
</tr>
</tbody>
</table>
Benefits

• Easier for technologists to contribute domain expertise via tests rather than policy language

• Empowers a wider range of people to contribute to the policymaking process

• Many others…
Questions?