Empirical Comparison of Language Workbenches

DSM Workshop, SPLASH, Indianapolis
27 October 2013

Steven Kelly
Contents

- Previous comparisons
- What to compare?
- How to compare?
- Experiment design
Previous comparisons

<table>
<thead>
<tr>
<th># of Tools</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Marttiin et al. 1993</td>
</tr>
<tr>
<td>5</td>
<td>Isazadeh 1997</td>
</tr>
<tr>
<td>2</td>
<td>Kelly and Rossi 1998</td>
</tr>
<tr>
<td>2</td>
<td>Kelly 2004</td>
</tr>
<tr>
<td>2</td>
<td>Pelechano et al. 2006</td>
</tr>
<tr>
<td>2</td>
<td>Ozgür 2007</td>
</tr>
<tr>
<td>10</td>
<td>MDD-TIF 2007</td>
</tr>
<tr>
<td>3</td>
<td>De Smedt 2011</td>
</tr>
<tr>
<td>5</td>
<td>El Kouhen et al. 2012</td>
</tr>
<tr>
<td>10</td>
<td>LWC 2011-2013</td>
</tr>
</tbody>
</table>
Previous comparison: MDD-TIF07

dsmforum.org/events/mdd-tif07/

Microsoft DSL Tools  Eclipse EMF/GMF/oAW

GME  MetaEdit+

© 2013 MetaCase
Previous comparison: El Kouhen et al. 2012

days to implement BPMN

- RSA: 12 days
- GME: 6 days
- MetaEdit+: 0.5 days
- Obeo: 5 days
- GMF: 25 days

tinyurl.com/gerard12
Previous comparison: LWC 2013


Total SLOC vs. Coverage

Below/right of curve = better

SLOC

3000

2500

2000

1500

1000

500

0

% Feature Coverage

20 30 40 50 60 70 80 90 100

Ensō  Más  SugarJ  Whole Platform  MPS  Onion  Rascal  Spofax  Xtext  MetaEdit+
What to compare?

<table>
<thead>
<tr>
<th></th>
<th>MDD-TIF</th>
<th>El Kouhen</th>
<th>LWC 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Language implementation</strong></td>
<td><strong>X</strong></td>
<td><strong>X</strong></td>
<td></td>
</tr>
<tr>
<td>Generators</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application implementation</td>
<td></td>
<td></td>
<td><strong>X</strong></td>
</tr>
<tr>
<td>Maintenance</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How to compare?

- **Feature coverage**
  - Subjective, influence on productivity unknown

- **Lines of code**
  + Objective, easy to measure
  - Not comparable across languages, styles

- **User satisfaction**
  - Strongly influenced by comfort zone

- **Time**
  + Objective, language-independent
  + Directly interesting variable

- **Cost**
  - Not comparable across countries, time, scale
Experiment design

- **Users:** Experienced: tool makers, consultants
  - Measure tool use, not learning curve
  - Spread cost of experiment

- **Tasks:** Focused on LWB use
  - Language implementation + generators

- **Compare:** Time
  - Video of completing each sub-task
  - Can’t reliably capture actual time, only time to repeat
  - Videos serve as advertising & tutorial for tool
    - Useful result, avoids users clicking and typing too fast

- **Objective, repeatable, checkable, economical**
  - Need to complement with broader + softer approaches
Thank you!

Questions?

info@metacase.com
www.metacase.com

USA:
MetaCase
5605 North MacArthur Blvd.
11th Floor, Irving, Texas 75038
Phone (972) 819-2039
Fax (480) 247-5501

Europe:
MetaCase
Ylistömäentie 31
FI-40500 Jyväskylä, Finland
Phone +358 14 641 000
Fax +358 420 648 606

"Very practical and highly recommended"
Computing Reviews

★★★★★ "Excellent", Amazon review