**ComQA:**

A **Community-sourced** Dataset for

Community QA websites reflect *information needs* of interest to *real people* that QA systems and search engines do not handle well.

**Complex Factoid Question Answering**

Beyond simple questions to those with *compositionality*, aggregation, *temporal* and numeric reasoning, and unanswerable questions.

with **Paraphrase Clusters**

Wide *lexical* & *syntactic variety* in question formulation

---

**Question topics**

- religion
- sports
- literature
- politics
- geography
- movies
- history
- language
- date
- person
- politician
- country
- university

**Answer types**

- location
- city
- mountain
- river
- village
- church
- castle
- temple
- monastery

**Results of baseline systems**

<table>
<thead>
<tr>
<th></th>
<th>WQ</th>
<th>Free917</th>
<th>ComQA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abujabal et al. (2017)</td>
<td>51.0</td>
<td>78.6</td>
<td>22.4</td>
</tr>
<tr>
<td>Bast &amp; Haussmann (2015)</td>
<td>49.4</td>
<td>76.4</td>
<td>21.6</td>
</tr>
<tr>
<td>Berant &amp; Liang (2015)</td>
<td>49.7</td>
<td>-</td>
<td>10.6</td>
</tr>
<tr>
<td>Berant et al. (2013)</td>
<td>35.7</td>
<td>62.0</td>
<td>12.0</td>
</tr>
<tr>
<td>Fader et al. (2013)</td>
<td>-</td>
<td>-</td>
<td>6.73</td>
</tr>
</tbody>
</table>

---

**Answer Annotations in ComQA**

- Semantically grounded answers: more informative & facilitate better evaluation.
- No commitment to specific answering resource → A dimension for innovation.

[qa.mpi-inf.mpg.de/comqa]