

## 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

Who was Britain's leader during WWI?

Who ran Britain during WWI?

A: [[https://en.wikipedia.org/wiki/h.\\_h.\\_asquith](https://en.wikipedia.org/wiki/h._h._asquith),  
[https://en.wikipedia.org/wiki/david\\_lloyd\\_george](https://en.wikipedia.org/wiki/david_lloyd_george)]

largest city located along the Nile river?

largest city by the Nile river?

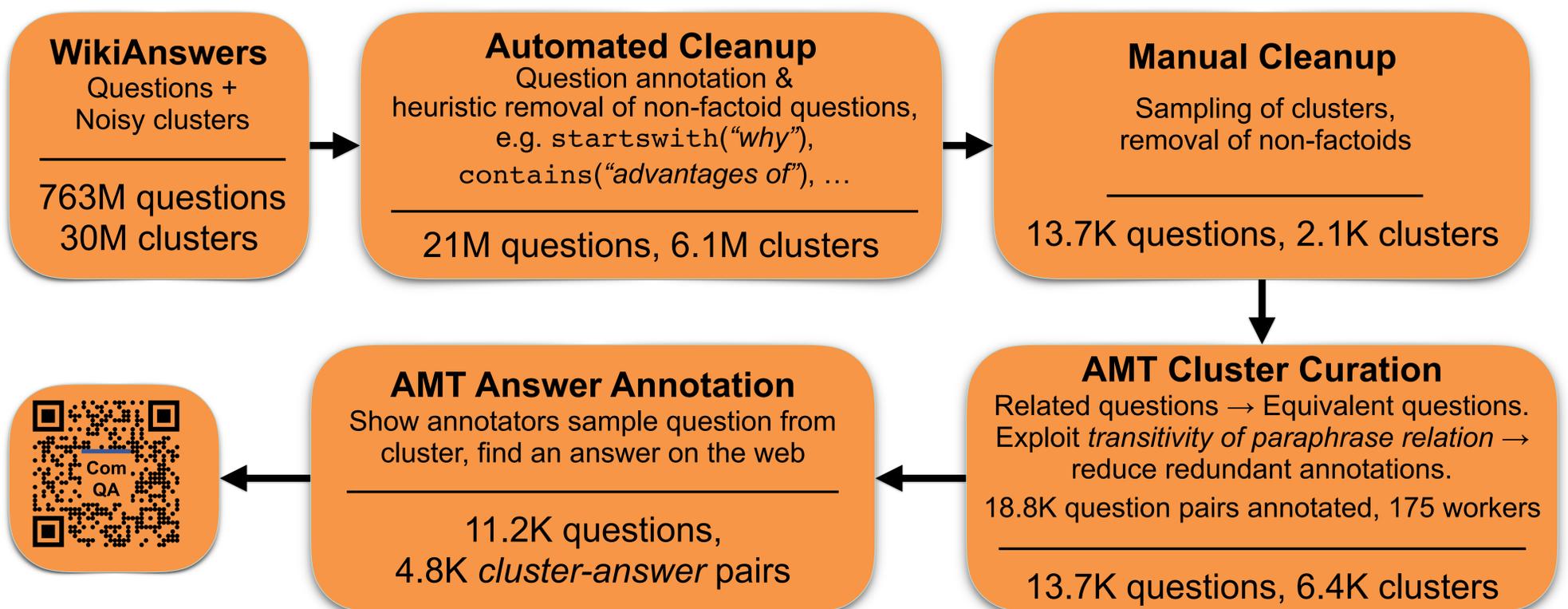
What is the largest city in Africa that is on the banks of the Nile river?

A: [<https://en.wikipedia.org/wiki/cairo>]

Who is the first human landed in Mars?

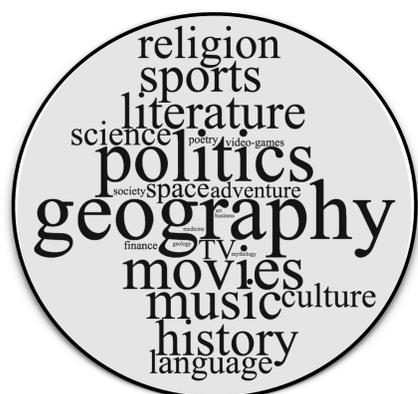
Who was the first human being on Mars?

first human in Mars? A: []



### Question topics

### Answer types



### Results of baseline systems

	WQ	Free917	ComQA
Abujabal et al. (2017)	51.0	78.6	22.4
Bast & Hausmann (2015)	49.4	76.4	21.6
Berant & Liang (2015)	49.7	-	10.6
Berant et al. (2013)	35.7	62.0	12.0
Fader et al. (2013)	-	-	6.73

### Answer Annotations in ComQA

- ★ Semantically grounded answers: more informative & facilitate better evaluation.
- ★ Answers: Wikipedia entities, TIMEX3 (temporal answers), SI (measurable quantities).
- ★ No commitment to specific answering resource → A dimension for innovation.