

The Semantic Elephant

Adrian Walker, Executable English LLC

For #data and #analytics, the elephant in the room is computational semantics in English. This is sometimes called the "semantic gap".

Practical all purpose natural language processing is hard to achieve, witness the spend on Watson and similar projects.

However there is a requirement -- question answering over relational databases -- that we can meet without the need to continually maintain dictionaries and grammars of English and all its sub jargons.

We can acquire knowledge in the form of Executable English syllogisms, like

some-person is a man

that-person is mortal

a-person became a British citizen on a-date1

that-person ceased to be a British citizen on a-date2

that-date2 is greater than that-date1

a-date3 is after that-date2

that-person ceased to be a British citizen between that-date1 and that-date3 [1]

Then, using the platform online at [2], we can run the syllogisms to answer questions, and to explain the answers. The platform automatically generates and runs networked SQL as needed.

Many question answering needs can only be specified using at least some requirements written in English. Yet there are maintainability reasons, as mentioned above, why direct specification in controlled English is avoided.

So, it may be useful to acknowledge the elephant in the room, and to run examples to see how open vocabulary Executable English can meet practical requirements for question answering over databases.

[1] https://www.executable-english.com/demo_agents/BritishNationalityAct1.agent

[2] www.executable-english.com