Interview: How can scientific methods provide guidance for Semantic Web Research and Development?

Kjetil Kjernsmo[1] and Ruben Verborgh[1]

[1] Department of Informatics, Norway
[2] Ghent University – iMinds, Belgium
kjetil@kjernsmo.net,ruben.verborgh@ugent.be

The organisers have asked Kjetil Kjernsmo to prepare a **interview** around his submission about an epistemological discussion around Semantic Web research. Together with one of the reviewers, Ruben Verborgh, he examined his contribution in this new format with success. The questions of the audience were answered during the succeeding discussion section.

Interview transcript

So, Kjetil, what's your confession? Like most people at this conference, I claim that my work advances the Semantic Web, but I can't prove that. In fact, I don't know any way to reach any conclusive result with respect to a larger Semantic Web vision, so in a way, all results are inconclusive.

We hope to advance the Semantic Web by doing research within the boundaries of the known. Isn't that all we can do? It seems hard to break out of the boundaries of the known, but no, I don't think it is good enough, and that we need to challenge this view, and that we should look to philosophy of science and natural science for inspiration. My fundamental assumption, which I hope you will agree with, is that science has had a remarkable rate of progress, and that we should adopt the practices that achieves this.

Now, we just end up testing one little piece at a time, based on what is already there, but the fact is that if we get to the visions that are presented as the Semantic Web, then it will be completely different in terms of workloads, data profiles, etc. And therefore, the central problem is that to really argue that we progress towards the Semantic Web, we have to test against a reality that doesn't exist yet. That makes the problem different from other sciences.

It seems that your background is not computer science then? Yes, I'm not a computer scientist. For my master's degree, I studied theoretical astrophysics. I only happened to bump into Dan Brickley for a side project in 1998, who went in to Chair the RDF working group the week after, and after I finished my master's, I've worked with Semantic Web in the computer industry. It is funny, I didn't work much as a physicist, but this training still makes me feel like a natural scientist. I've also read a lot of history and philosophy of science. I think that we might find a key to speed up progress if we start discussing epistemology.

What does "epistemology" mean? Naively, epistemology is the study of what it really means to say that you know something. So, how do you really know that a triple store performs better than another? How do you know that a certain ontology engineering approach will forward the Semantic Web and not hinder it?

I understand your frustration with the lack of progress and the lack of methodological clarity, but I'd like to understand why you're concerned with this. Ok, I'll will introduce and example from modern astronomical history. I studied quasars, which are active central parts of very remote galaxies. Rather spectacular and complex stuff, there was initially a lot of confusion, but many possible directions were quickly terminated. Like, it was quickly clear that they were remote objects, just not how insanely remote they were. After a couple of years, people were finding the spectra, determining redshift, etc. Very cleverly, but with a clear methodological direction, after a few years, the picture of this very remote, complex thing became quite clear, and has largely stood the test of time. Now, the big difference is that with quasars, we were discovering a world that existed, but that we didn't know about, but we quickly figured out how to understand. With the Semantic Web, there's a world that we're creating, but we don't know how to understand how to get there. So, if you accept the premise that science shows a remarkable rate of progress, then I would like to discuss and understand what it is about science that does this.

But isn't the Semantic Web progressing equally fast? First of all, what do we mean by the Semantic Web?

A machine-readable Web of data... I think it has to be much broader defined, for one thing, it is also machine writeable. In the keynote at ISWC 2011[3], Frank van Harmelen discussed the the heterogeneity problem and argued that the solutions are not so much technical, but social, economical and cultural, and a definition of the Semantic Web must reflect that. I've been around for 17 years, and I don't think it is progressing very fast. If you look at job postings, you'll find that the number of jobs containing Semantic Web is varying but low, and orders of magnitude lower than e.g. MongoDB. Also, Google Trends shows a steady decline. And we still spam W3C mailing lists with CfPs, even though if we had the Semantic Web, it should be easy to create a system to match a paper to a venue. I liken the Semantic Web to a flying car: People had visions about flying cars and how they would impact society. Guess what: flying cars exist, they have been built. But there are 5 of them. Or maybe 10. But clearly, their mere existence does not give them the impact they had in those visions.

There has been some papers on research methodology, for example, Avi and Natasha have their "Is This Really Science? The Semantic Webber's Guide to Evaluating Research Contributions" [1], doesn't that provide enough guidance? It is an excellent read, but it largely adopts a falsificationist view of science,

[3] http://www.cs.vu.nl/ frankh/spool/ISWC2011Keynote/

as it was argued by Karl Popper. He said that one should seek to formulate bold conjectures, so that if we're wrong, then a test should easily show that. So, considering the epistemology, the things that were knowable are the things that you can formulate a falsifiable hypothesis for. The funny thing is that falsificationism worked really well for me when I studied quasars. There are two large theories that I worked within, Theory of General Relativity and Quasar Theory. It was easy, even for a master's student, to come up with conjectures that were falsifiable, not only for my own work, but observations done within the framework I developed could potentially brings parts of the large Quasar Theory down with it. That's of course a long shot, but it is really cool to know that when you're a student.

Why can't we just apply falsification to the Semantic Web? That's what researchers do—and successfully, it seems? For the reasons that falsification has been criticized for by more recent philosophers. I would recommend "What is this thing called Science" by A.F. Chalmers [2]. The TL;DR version is "science is pretty darn hard, especially when doing research". By that, I mean that you can always construct a philosophy of science around conjectures like "all swans are white", but as researchers our conjectures are never that simple. For example, you can never know if it is really your hypothesis that is wrong, or the evaluation. So, it might both happen that your hypothesis is really true, but you rejected it due to a faulty evaluation methodology, or your hypothesis was false, but you failed to reject it, because the evaluation was not good enough.

Do you have an example of this? There was a good talk at ISWC last year, where the author analyzed several SPARQL benchmarks, and showed how each of them were unsuited for important real-world considerations. However, the problem is that if you applied the same methodology to his own study, that too would have fallen. So, to gain any knowledge, you always have to question your methodology. Moreover, it could be that there are auxiliary hypotheses that you simply accepted, it could be that your hypothesis had not been rejected if the triple store data structure had been a trie and not a B-tree.

So falsification does not work in all cases? Some philosophers have trolled the scientific community and said stuff like "falsificationism would destroy all of science". For example Tycho Brahe correctly deduced that Copernican theory predicted an observable property, which he, despite his instruments being orders of magnitude better than his predecessors, could not see. So if you really adopted falsificationism, you'd probably reject prematurely in many cases. Besides, what if falsificationism was applied to the Semantic Web visions, wouldn't we have rejected the whole thing by now, stopped our conferences and gone on to do something else? So, no, it is not just that falsificationism doesn't work well, it is that we wouldn't really have gotten started, and isn't really what scientists, historically, at least, have been doing.

At ISWC2013, you presented a paper that attacked benchmarking [3], and proposed to use statistical design of experiments instead. Is that part of that picture? Yeah, I wanted to illustrate a methodology where you could actually test the evaluation to see if the methodology stands up to scrutiny.

But the flaw here seems to be that you would test the methodology, but you need another test of the test methodology, to evaluate that methodology, and so a test of that methodology again, ad infinitum... Yeah, that's the kind of philosophical problem I'm struggling with. There's a recent direction in the philosophy of science called the "New Experimentalism" that focuses on experimentation, instrumentation, laboratory practices, that goes further in this direction, where it is not sufficient to test a hypothesis, you need to prove that the test is severe, and they go on to formalize what severe means. I have not yet appreciated the broad relevance of this philosophy, but the people working on this are still alive, and it seems promising.

As you talk about hypotheses and theory, can you formulate a theory of the semantic web, like you have the quasar theory? I think it is perhaps one of the most important exercises we can do, but I'm not confident enough to attempt that, it must necessarily be a community effort.

Ok, but can you explain what is meant by "hypothesis" and "theory", then? Maria-Esther forwarded Oxford definitions, but I don't think you'd find consensus about those if you asked philosophers, it is remarkable. Superficially, I think hypotheses are targeted conjectures, easy come, easy go things, things that you try to make testable in everyday research. A theory is a collection of hypotheses that have been tested.

A collection, can you attempt a more stringent definition? Well, I can try a definition that I personally find good, but that doesn't mean it is well founded in philosophy of science. It would be "A theory is a coherent set of hypotheses...".

Coherence, what would that be in this context? Coherence places a burden on the researcher when considering the big picture: If my work finds that I get a performance benefit in spite of a longer connection time, it will be incoherent with a hypothesis that connection time must always be kept to a minimum. There are some major schools of thought called coherentism, which requires that all hypotheses are not only consistent, but also meet the requirements of e.g. Occam's razor. To me, pure coherentism appears to have some merits when dealing with large, theoretical frameworks, such as string theory, or indeed much of the work on logic that this community does. A good thing to say about it would be that it might have many points of attack, and a successful attack could bring the whole theory down.

What do you mean by "attack", is that a to test the theory against reality? Yeah, so my full definition would be something like "A theory is a coherent set of hypotheses that are held as true after having been subjected to vigorous testing by a scientific community".

So this brings up social aspects of science. This is something the physicist and philosopher Thomas Kuhn emphasized, right? Yes, but I'm not a big fan of Kuhn. He introduced some interesting concepts, like "normal science", "scientific revolutions", "paradigms", and the emphasis on science as a social process. So, yes, with this definition makes it important that the community is capable of vigorous testing. I think he also makes an important point in that he thinks philosophy of science must be based on history of science, and in that way, he differs from Popper. The problem is, he got significant parts of his history wrong, and for that reason, I reject most of his philosophy.

You opened by saying that every result is inconclusive. I suppose that that can be said for anything, also in physics. For example, Newtonian mechanics is inconclusive in retrospect, since it has been replaced by relativity. So, what is different in our field? Yeah, that's a fair point to argue, but our case is very different, because our hypotheses are tested against a reality that doesn't exist, as the Semantic Web doesn't exist yet. Once the Semantic Web exists, it will be very different from what we have today, like if everyone had a flying car, we'd have a completely different set of traffic regulations. So, you can take the exception, and say that this is how Computer Science operates, and there's nothing you can do about it. If you do that, how can then this field achieve the rate of progress that natural sciences has? I'm arguing that this isn't good enough, and that's why I initiate this discussion. In all fairness, I think that if you claim that you are forwarding the Semantic Web (and I suppose you do if you're at this conference), you have the burden of proving that you do. But that can only be done if we devise a way to test our hypotheses against a more clearly defined Semantic Web Theory.

So, to round off, what do you suggest we do? First, we need to acknowledge there's a problem, that just testing pieces of the big picture isn't good enough. Secondly, we have to start formulating theories. Then, we have to increase the prestige of empirical methods. For example, can we formulate our tests as "severe" with the "New Experimentalism"? And, finally I think we should spend more time to sit down and hack and bring the code to the people.

References

1. Abraham Bernstein and Natasha Noy. Is this really science? the semantic webber's guide to evaluating research contributions. Technical report, 2014.
2. Alan F Chalmers. *What is this thing called science?* Hackett Publishing, 2013.
3. Kjetil Kjernsmo and John S Tyssedal. Introducing statistical design of experiments to sparql endpoint evaluation. In *The Semantic Web–ISWC 2013*, pages 360–375. Springer, 2013.