# Good Scientists Pimp there Research (was: Damn, I'm boring...)

Egon Willighagen

## Citation

## Keywords

Chemistry

## Abstract

Define good. Let me say that up front. Good scientists, that is, if you say successful researchers are good scientists, secure good funding. Getting good funding requires doing the most relevant research (define relevant). Or, to put it bluntly, being a successful researcher requires to pimp your research. Doing boring research is nice for you, good for a Nobel prize if it turns out to have a cool spin off, but doesn't buy you research success.

## Copyright

# chem-bla-ics

Define good. Let me say that up front. Good scientists, that is, if you say successful researchers are good scientists, secure good funding. Getting good funding requires doing the most relevant research (define relevant). Or, to put it bluntly, being a successful researcher requires to pimp your research. Doing boring research is nice for you, good for a Nobel prize if it turns out to have a cool spin off, but doesn't buy you research success.

And, boy, am I boring or what? Really, Peter is right... I am boring. He is also right that doing quality control doesn't result in research output. He writes:

> *But it's essential for modern knowledge-driven science. The chemical software and data industry has no tradition of quality. I've known it for 30 years and I've never seen a commercial company output quality metrics. I have never seen a commercial company publish results of roundtripping. That's another really boring and apparently pointless operation where you take a file A, convert it to B and then convert it back to A'. What's the point? Well A and A' should be the same. With most commercial software you get loss. If you are lucky it's only whitespace. But it's more likely to be hydrogens or charges or whatever.*

There is a huge gap between *should* and *is*. This is partly caused that it is much easier to throw garbage at some piece of code, than to start an Ugi reaction on something which is naturally broken. Take a moment to think about that. *Garbage in, garbage out* is still common in information technology, but I have never heard of the saying *oil in, oil out*.

**Scientists do not have the ability to recognize garbage (when it comes to data)!** It's statistics, that other boring stuff I do. Even worse, it's the programmers fault if someone throws garbage at my software and it doesn't return 42.

Unit testing is a poor mans approach to handle garbage. As are diagnostics tools, like finding the difference between IChemObjects (and #2 ). Unit tests define common *garbage in* (of course, also *valid input in*, but that does not make it less boring) and tests that the method does not freak out; that the tested method gives the user some friendly message like *"Hi there! I'm your friendly computer program. I am sorry to inform you that it's garbage you just passed along. Please clean it up first."*

Anyway, back to my boring research...