



I Choose You, Fact 44!

**Understanding what is there is harder than
just writing something new**



Summary of Fact 44

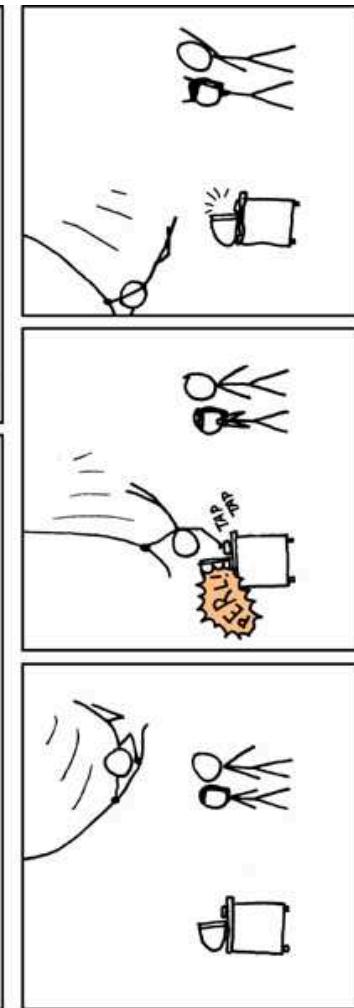
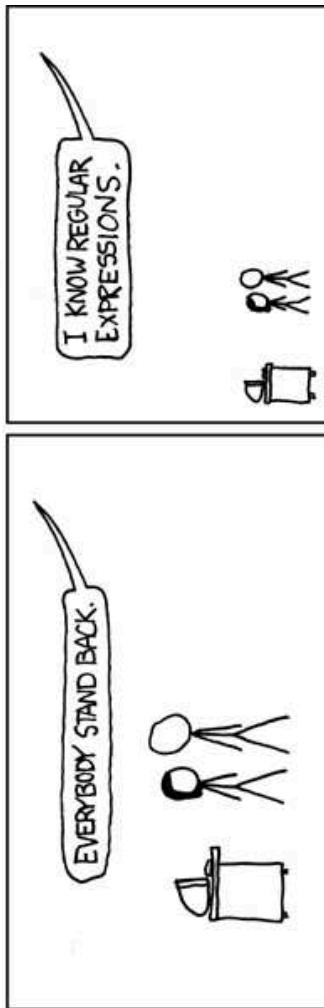
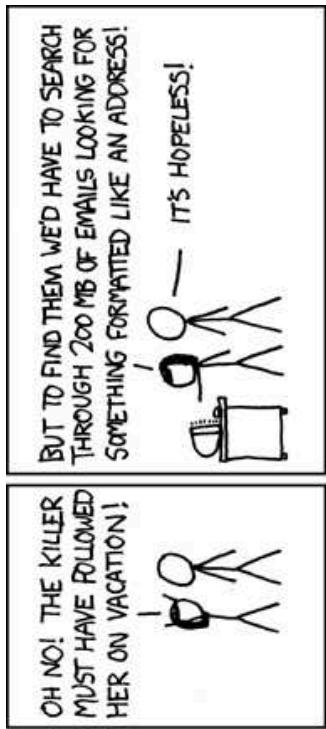
- Programmes are written once, read several times
- Every reader has to re-understand
 - intent of the code
 - what the code actually does
- Maintenance documentation missing almost always
 - reasons *against* certain design decisions not clear
 - “Let’s make the same mistake (over and over) again.”
- Have you ever read a piece of code and thought “why the hell is this working?” or “what on Earth is this doing?”
- Result: high costs for maintenance



Connection to Own Work (1)

- Next Generation Sequencing Data
 - paired reads = sequences belonging to each other
 - two files, pair = entry of one file + entry of other file (same IDs)
 - processing can lead to losing one partner of a pair
 - using sequences to build a genome
 - paired data must be complete
 - need way to remove singlettons, so only pairs are left
 - select_paired.pl does that
 - not... because our IDs do not match “^ (> . *) \ / \ d\$“
 - ☺

Connection to Own Work (2)



Source: <https://xkcd.com/208/>
Author: Randall Munroe

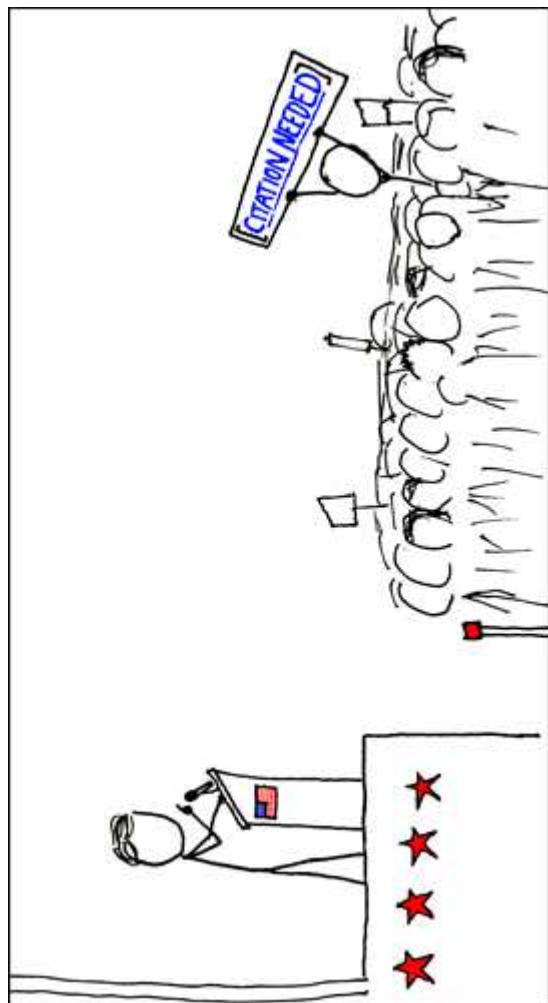


Connection to Own Work (3)



Fact 44 is true[citation needed]

- I experience this *very often*.
- Actually, every time I look at other people's code.
- Or even my own code from years ago
 - “OMFG! I cannot believe, I did it *that way* back then.”



Source: <https://xkcd.com/285/>

Author: Randall Munroe



What I learned from this

- There is something called ‚maintenance documentation‘.
- I had this idea/notion of something like that in my head, but thought it was just one of my unrealistic perfectionist ideas.
- But: There is a good chance that maintenance documentation would be actually useful and beneficial.
- I want to encourage everyone to create maintenance documentation. This needs to be the standard.