

Technical Debt

Colin Gillespie (@csgillespie)

colin@jumpingrivers.com

O'REILLY®



Efficient R Programming

A PRACTICAL GUIDE TO SMARTER PROGRAMMING

Colin Gillespie & Robin Lovelace

About Me

Co-Founder: Jumping Rivers



 @jumping_uk

 @csgillespie

jumpingrivers.com

Jumping Rivers

Data Science Consultancy and Training

✉ hello@jumpingrivers.com 🐦 @jumping_uk

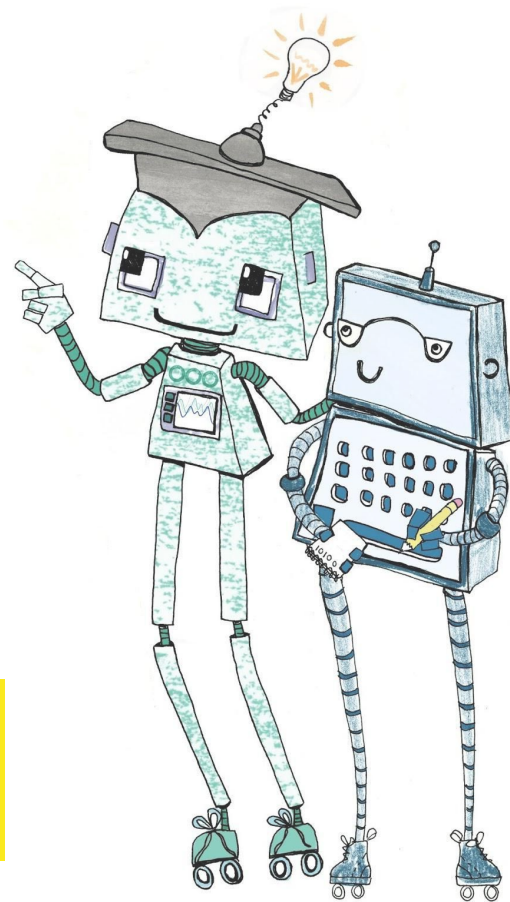
All things data science

Training

Machine learning, DevOps, infrastructure

Managed RStudio services

Shiny development and deployment





What is technical debt?
...cost of additional rework



What is technical debt?

...cost of additional rework
caused by choosing an easy
(limited) solution now



What is technical debt?

...cost of additional rework
caused by choosing an easy
(limited) solution now
instead of using a better
approach that would take
longer

wikipedia

Where does tech debt come from

Business pressures*

A

KA blame your boss



My boss

*This is one of many reasons



We are not software engineers*

Smaller teams

Not building large pieces of softwares

In general*



Data Scientists + Tech debt

One-off pieces of work



Data Scientists + Tech debt

One-off pieces of work

Statistical analysis



Data Scientists + Tech debt

One-off pieces of work

Statistical analysis

Relatively simple pieces of software



Tech Debt Smells*

You are frightened to change things

Yes, I am mixing metaphors*
Blame Mohammed for only given me 10 mins



Tech Debt Smells*

You are frightened to change things

Handing over projects is hard

Yes, I am mixing metaphors*

Blame Mohammed for only given me 10 mins

Tech Debt Smells*

You are frightened to change things

Handing over projects is hard

Low probability code will run

Yes, I am mixing metaphors*

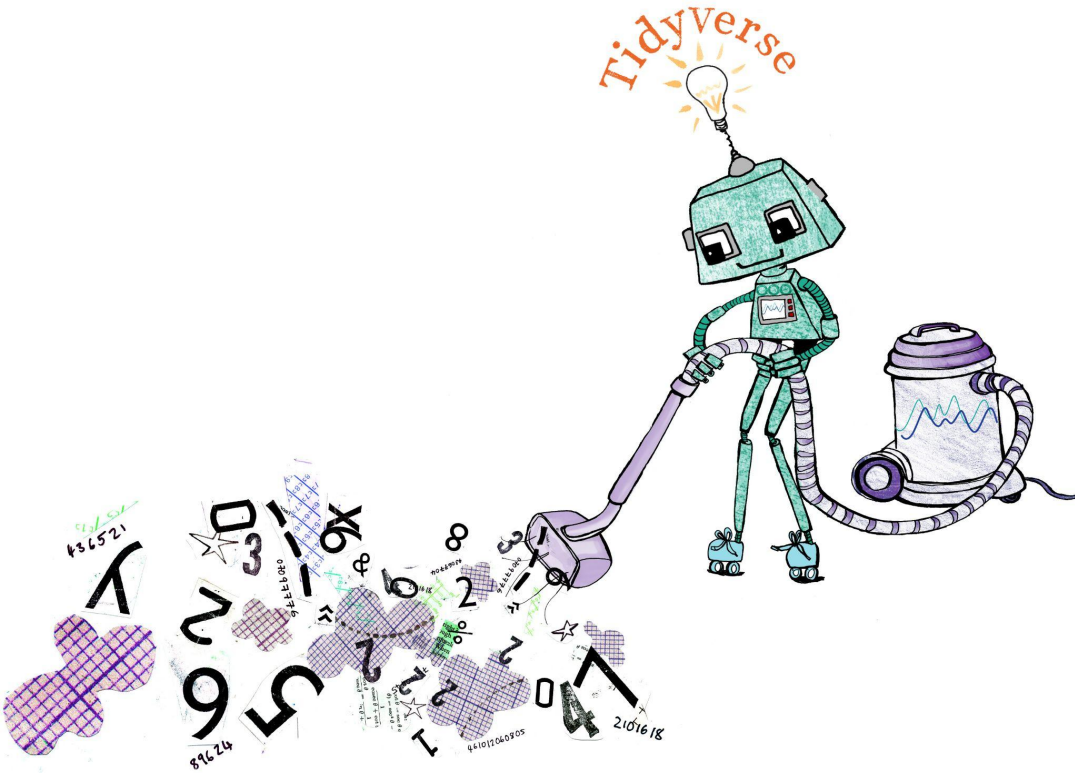
Blame Mohammed for only given me 10 mins

What to do?

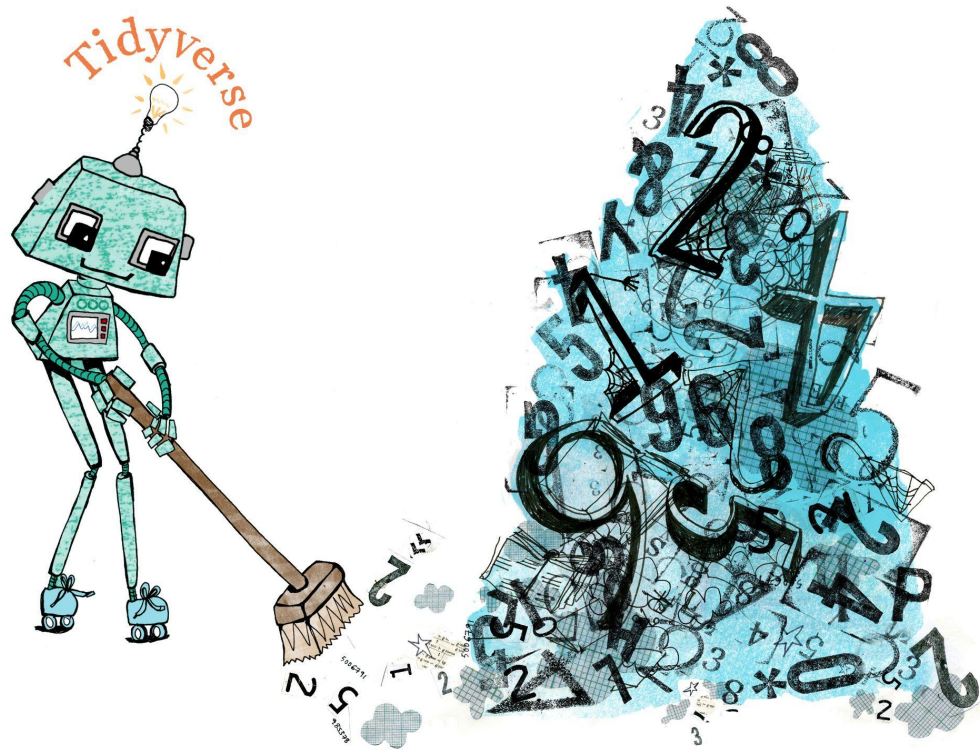
Fix the little things

Whenever you make change

Fix something small



Whenever you make change*



Update the README

Rename some variables

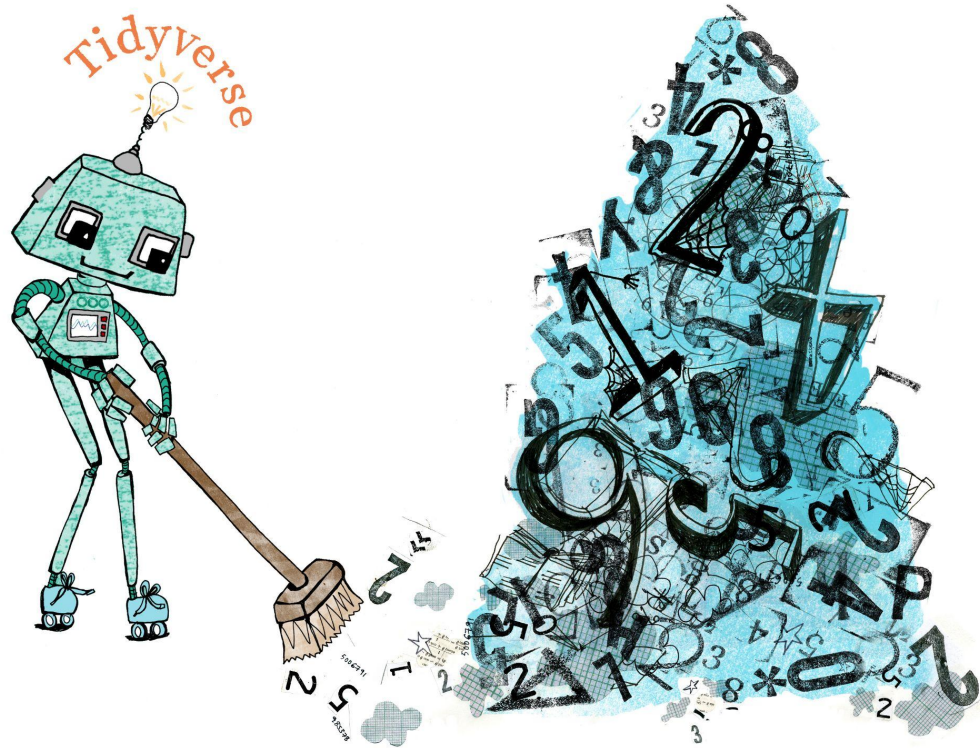
Fix the small things, Kent Beck; NSFW

Rename some variables

Add a test

Fix the small things, Kent Beck; NSFW

Whenever you make change*



Update the README

Rename some variables

Rename/split out a file

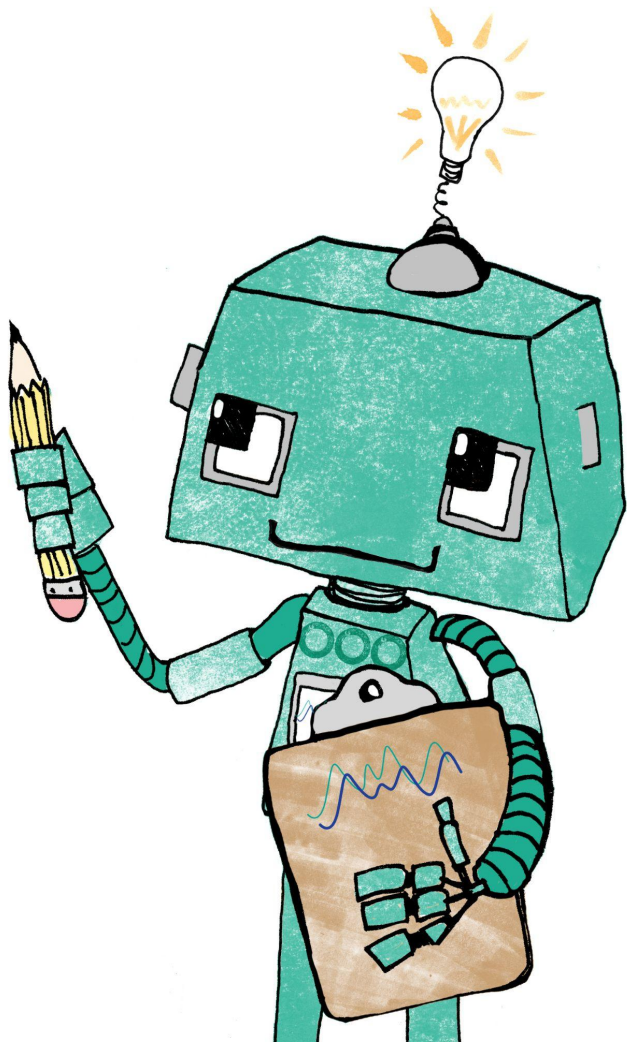
Add a test

Refactor a function

Fix the small things, Kent Beck; NSFW

Now the hard part





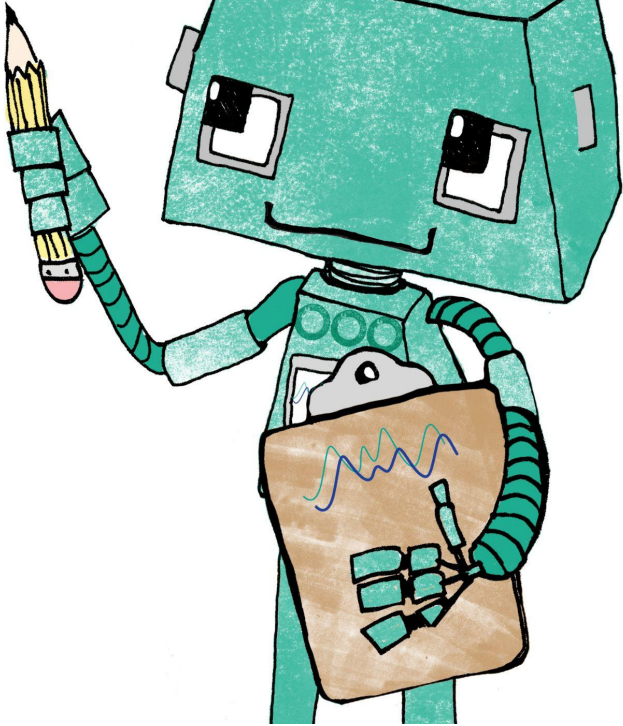
Agree **Red** Lines

R packages: 0 Notes/Warnings

{lintr}

CODEOWNERS

Code Review + Continuous Integration

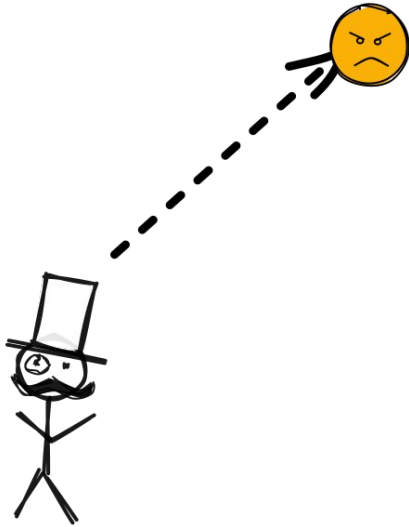


Data Scientist

Git



main



Data Scientist

Git

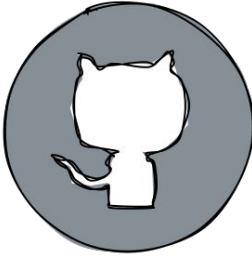
Applications



main



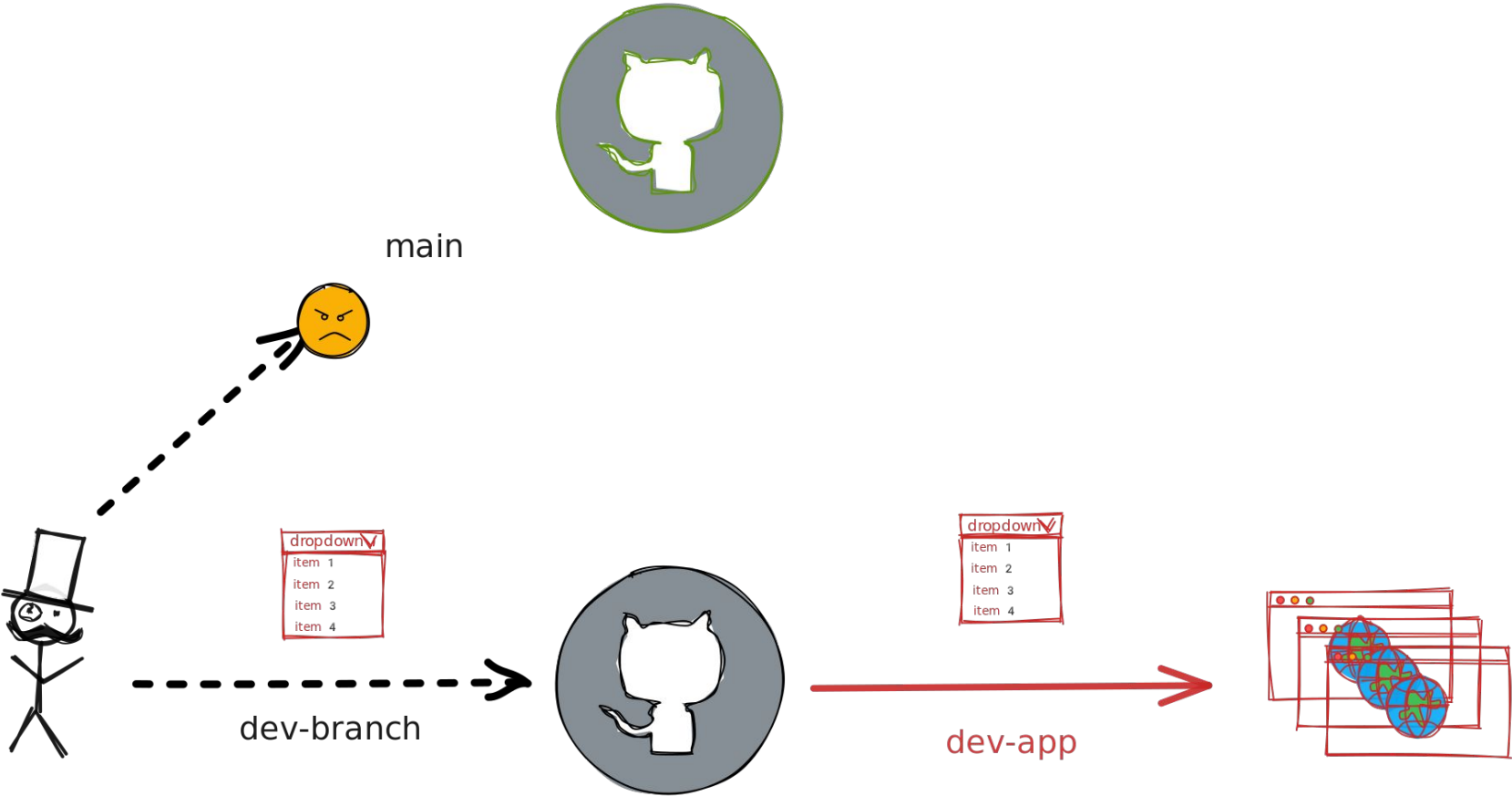
dev-branch



Data Scientist

Git

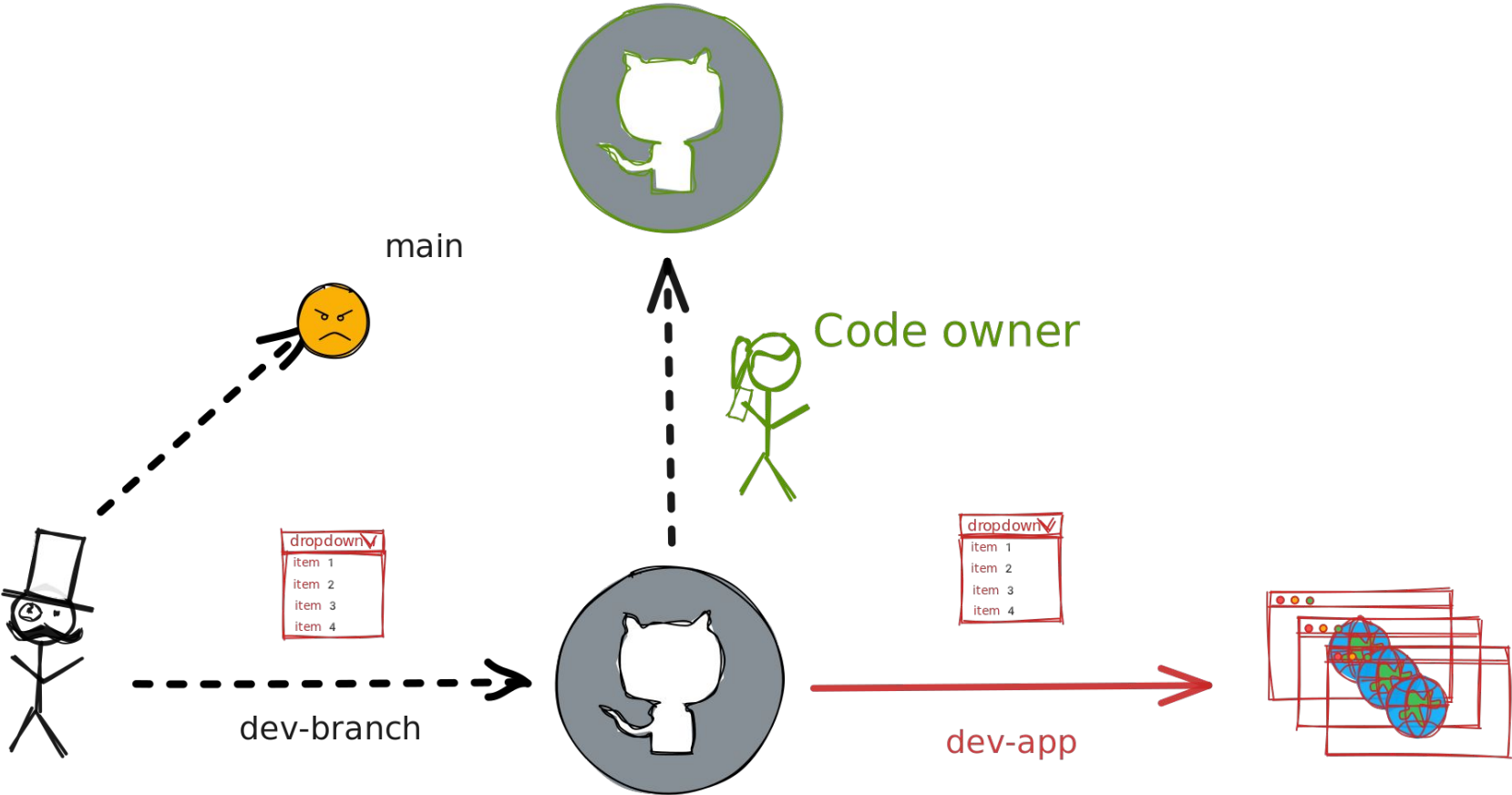
Applications



Data Scientist

Git

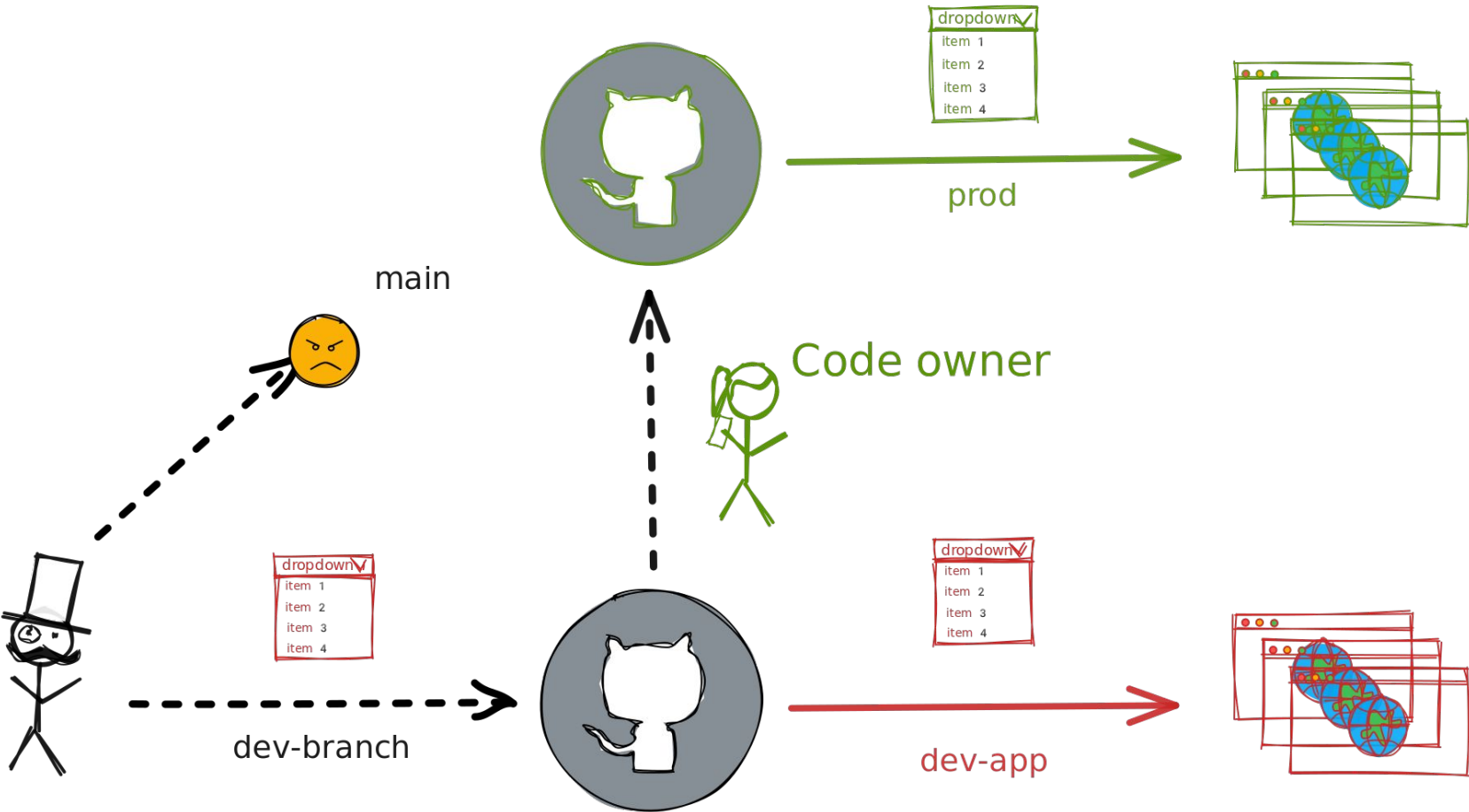
Applications



Data Scientist

Git

Applications

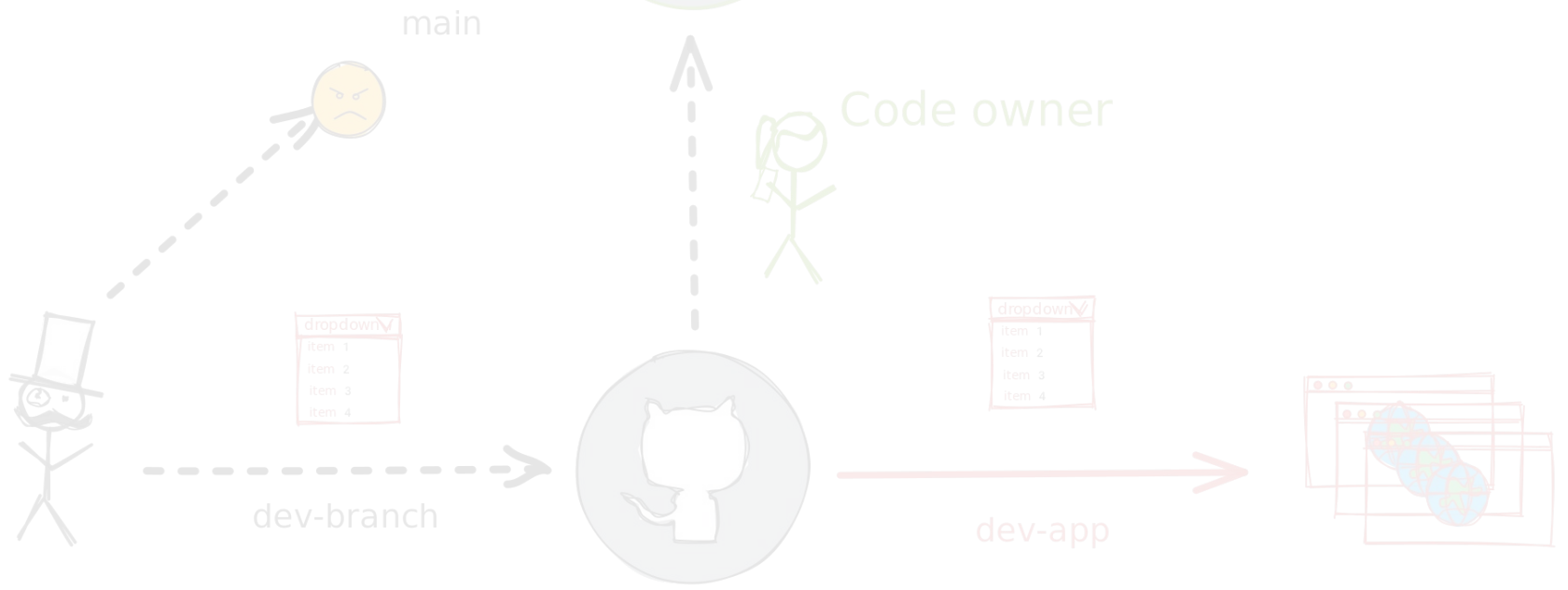


Data Scientist

Git
Before you say what is/isn't required

Applications

How important is your work?

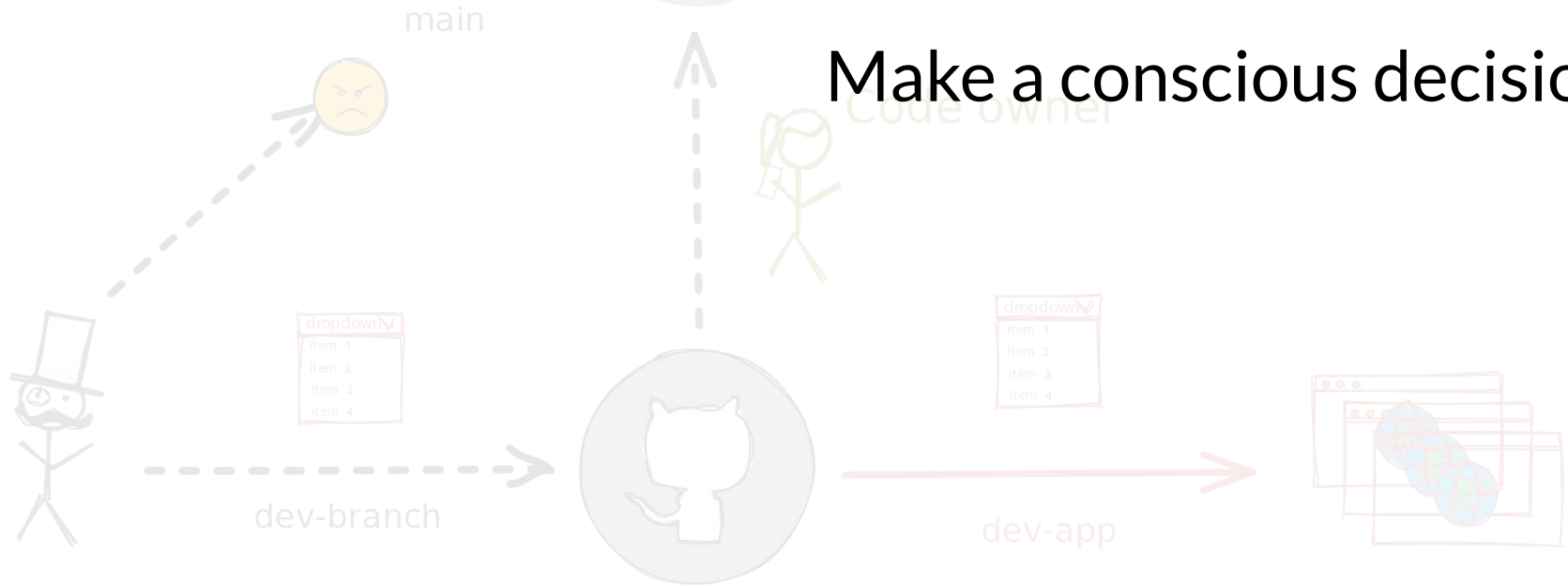


Data Scientist

Git
Before you say what is/isn't required
Applications

How important is your work?

Make a conscious decision



Data Scientist

Git
Before you say what is/isn't required

Applications

How important is your work?

Make a conscious decision

It might be OK to be sub-optimal*





Summary

Data science tech debt is different; but still important

Summary

Data science tech debt is different; but still important

Fix little things

Use Git / CI effectively

Summary

Data science tech debt is different; but still important

Fix little things

Use Git / CI effectively

Blame the boss

Questions?