CODING CHALLENGE AND EFFECTIVE TEAM WORKING

Mark Dawson, Ed Bennett, Michele Mesiti, Colin Sauzé

AIMLAC CDT Induction, 2020-09-22

WHY ARE WE DOING THIS?

- Gain experience of
 - Remote working and collaboration
 - Applying machine learning / big data to real(ish) problems
 - Technical skills
 - Programming
 - Data processing
 - Machine learning
 - Cloud Computing
 - Software Engineering
 - Version control
 - Testing/Continuous Integration
- Something to show/talk about to prospective employers
- Prepare you for your placements

WHAT YOU'LL BE DOING

- Project running until May
 - Presented at end of year event in June
- A few hours per week of work
- Regular meetings with a "customer"
- Help sessions with RSEs

ZOOM THINGS

- Please mute when you're not speaking
- If you're comfortable having a camera on, it's nice to see who we're talking to
- The Participants window has options "raise hand", "yes", "no", etc. that we will use
 - Select "Yes" now
- 3 breakout rooms
 - Each one gets an Etherpad: https://etherpad.wikimedia.org/p/aimlac-2020-TEAMNUMBER
 - We'll let you know when to go and come back

INTRODUCTIONS

- Name
- PhD topic
- What were you doing this time last year?

SCENARIO

- You are a small software team specialising in bespoke machine learning products
- You've been hired by an (imaginary) customer
- They wish to investigate automating business decisions with Artificial Intelligence, Machine Learning and/or Advanced Computing

WHAT MAKES TEAMS CHALLENGING?

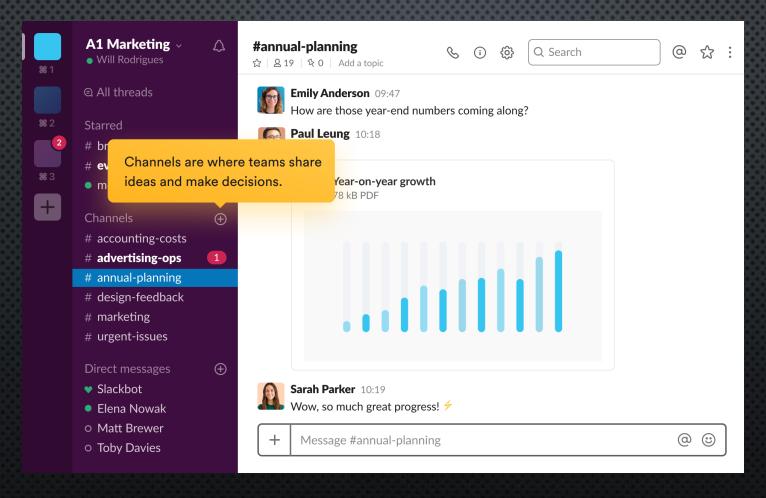
- What difficulties can you see with delivering software as a small team, responding to customer needs?
- Discuss in your breakout rooms
- Write thoughts down on the Etherpad
- 10 minutes

HOW CAN TEAMS?

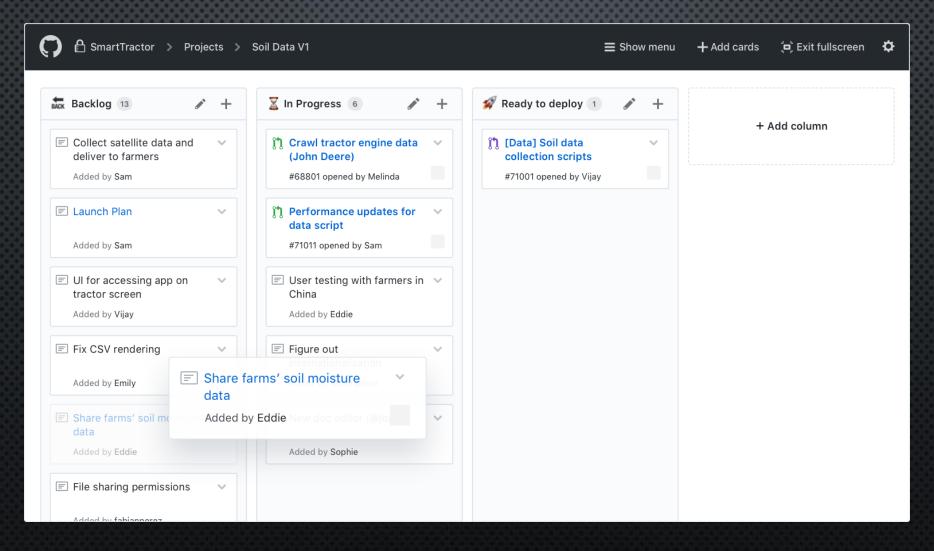
- Can you imagine some ways of working that might resolve these problems?
- Discuss in your teams
- Agree a series of short bullet points (max 7) describing the process you might use
- 10 minutes

TOOLS!

Slack

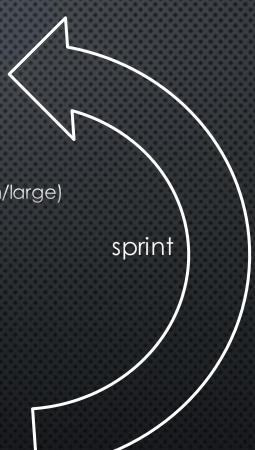


7 GitHub



SAMPLE PROCESS

- Meet the customer
- Hold a planning meeting
 - Agree tasks
 - Add tasks to your task list as "waiting"
 - Estimate and note task length (small/medium/large)
 - Reflect on the last iteration
- Choose a task (waiting or review)
- Self-assign the task
- If "waiting" then move task to "active"
- Complete the task
- Move the task onwards
- Meet with your technical contact
- Have a "daily standup" at agreed intervals



STANDUP

- Three questions:
 - What have you achieved since last standup?
 - What will you work on for next standup?
 - What issues might you anticipate?

PROCESS DEVELOPMENT

- The ideal process is different for every team and organization
- Make looking critically at your process and developing it part of your process

EFFECTIVE SOFTWARE DEVELOPMENT

Mark Dawson, Ed Bennett, Michele Mesiti, Colin Sauzé

AIMLAC CDT Induction, 2020-09-23

What are the challenges to writing software?

- Writing software is about telling computers what to do
- Why might this be challenging for humans to do effectively?
- Discuss in your teams
- Write thoughts down on the etherpad:

 https://etherpad.wikimedia.org/p/aimlac-2020-team1
 https://etherpad.wikimedia.org/p/aimlac-2020-team2
 https://etherpad.wikimedia.org/p/aimlac-2020-team3
- 10 minutes

HOW CAN WE MITIGATE THESE DIFFICULTIES?

- What techniques, methods, language features and tools would help mitigate these difficulties?
- Discuss in your teams
- Write thoughts down on the etherpad:

 https://etherpad.wikimedia.org/p/aimlac-2020-team1
 https://etherpad.wikimedia.org/p/aimlac-2020-team2
 https://etherpad.wikimedia.org/p/aimlac-2020-team3
- 15 minutes

SYNTAX HIGHLIGHTING

```
1 context.py 2<sup>3</sup> utils.py •
   class ContextPopException(Exception):
        "pop() has been called more times than push()"
        pass
   class ContextDict(dict):
        def __init__(self, context, *args, **kwargs):
            super(ContextDict, self).__init__(*args, **kwargs)
            context.dicts.append(self)
            self.context = context
        def __enter__(self):
            return self
        def __exit__(self, *args, **kwargs):
            self.context.pop()
context.py [python][unix→utf-8]
                                                          L17/267:C0 4%
```

AUTO-COMPLETE

```
1 import numpy as np
   import matplotlib.pyplot as plt
   import pandas as pd
  x = np.arange(-np.pi,np.pi, 0.05)
  y = np.sin(x)
   plt.plot(x,y)
   plt.title('plot with numpy data')
   df = pd.DataFrame(y, index=x)
   df.plot(title='plot with pandas data')
12
13 np.
14
      abs
                        instance
      absolute
      absolute_import
      add
      add_docstring
      add_newdoc
      add newdoc ufunc function
      add_newdocs
      alen
      all
```

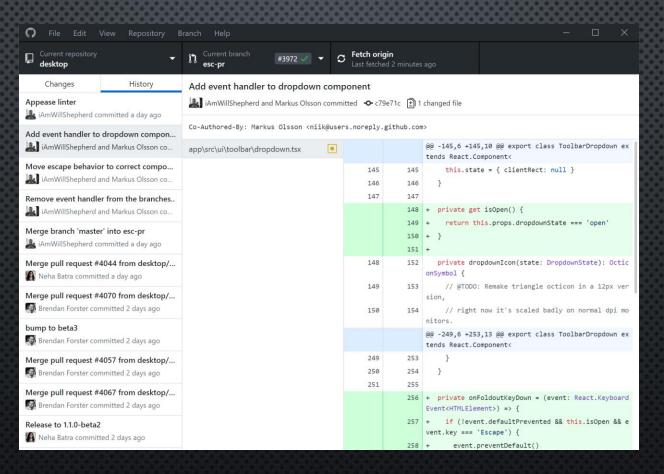
Inline Documentation

```
def StormCastle(castleName, castleLocation, totalTroops, veapons):
     :param string castleLocation: The location of the castle.
     print "Storming: " + castleName + " at " + castleLocation
     print "Total Troops: " + str(totalTroops)
     print "Weapons: "
     for weapon in veapons:
         print " * " + weapon
 StormCastle()

☐ IronPythonTest

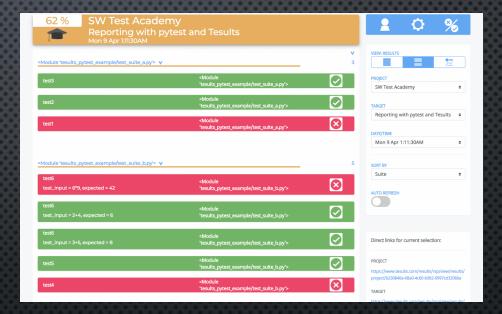
 Storm a castle anywhere in the world.
```

Version Control

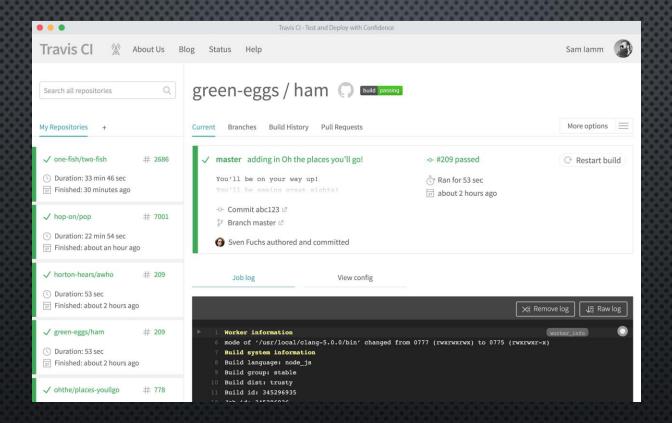


AUTOMATED TESTING

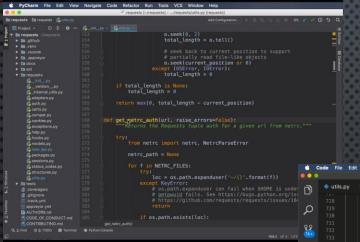
```
$ py.test -v --tb=no
                      ----- test session starts -----
platform darwin -- Python 2.7.2 -- py-1.4.20 -- pytest-2.5.2 -- /Users/John/.pyt
honbrew/pythons/Python-2.7.2/bin/python
collected 27 items
ctest/test:0: test1 PASSED
ctest/test:12: test2
ctest/test:16: test1
ctest/test:0: test3 PASSED
ctest/test:0: test1 PASSED
ctest/test:0: test3 SKIPPED
ctest/test:53: test2 F
ctest/test:61: test1
ctest/test:0: test1 PASSED
ctest/test:0: test2 PASSED
ctest/test:0: test1 PASSED
ctest/test:98: test_assert_str FAILED
ctest/test:103: test_assert_equal FAILED
ctest/test:108: test_assert_not_equal FAILED
ctest/test:114: test_assert_null FAILED
ctest/test:0: test_assert_not_null_const PASSED
ctest/test:123: test_assert_not_null FAILE
ctest/test:128: test_assert_true FAILED
ctest/test:133: test_assert_false FAILED
ctest/test:0: test_skip SKIPPED
ctest/test:141: test_assert_fail FAILED
ctest/test:0: test_null_null PASSED
ctest/test:150: test_null_string FAILED
ctest/test:154: test_string_null FAILED
ctest/test:0: test_string_diff_ptrs PASSED
ctest/test:0: test_large_numbers PASSED
ctest/test:0: test_ctest_err FAILED
      ======== 15 failed, 10 passed, 2 skipped in 0.09 seconds ==
```



CONTINUOUS INTEGRATION

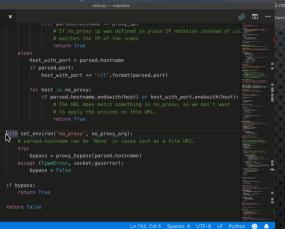


AUTO-REFACTORING



Change signature

Extract function



Rename file/module

PyCharm File Edit View Navigate Code Refactor Run Tools VCS Window

3o to File ☆#O

▼ I requests

► b _appveyo

__version__.py
__internal_utils.py

auth.py
certs.py
compat.py
cookles.py
exceptions.py
help.py
hooks.py
models.py
packages.py
sessions.py

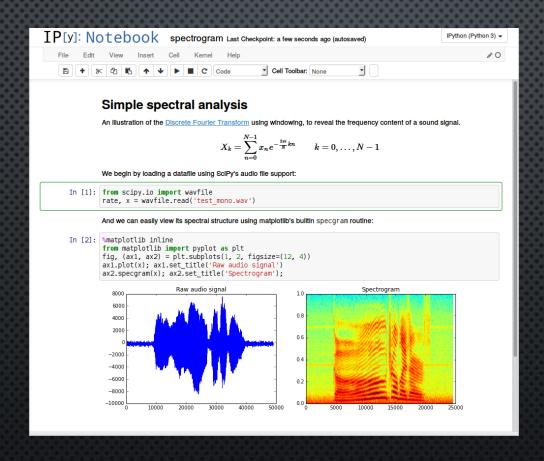
SYNTAX CHECKERS

```
?# encoding: utf-8
!import datetime
from south.db import db
from south.v2 import DataMigration
from django.db import models
!class Migration(DataMigration):
    def forwards(self, orm):
         pass
    def backwards(self, orm):
         "Write your backwards methods here."
             'foo': {'object name': 'bananas!'},
             'foobaz': ('django.db.models.fields.CharField', [], {'max length': '100'}),
             'foobaz': ('django.db.models.fields.AutoField', [], {'primary_key':, 'True'}),
             'foobaz': ('django.db.models.fields.CharField', [], {'max length': '200'}),
             'foobaz': ('django.db.models.fields.BooleanField', [], {'default': 'False'}),
```

AUTO-FORMATTERS

```
example.py
   x = { 'a':37, 'b':42,}
  'c':927}
   y = 'hello ''world'
6 z = 'hello '+'world'
   a = 'hello {}'.format('world')
   class foo ( object ):
    def f (self ):
     return 37*-+2
11 def g(self, x,y=42):
     return y
   def f ( a):
14 return 37+-+a[42-x : y**3]
15
```

LITERATE PROGRAMMING



What does good code look like?

- Each example has many positives and negatives
- Discuss in your teams. What are the good points? What are the bad ones?
- Mhàs
- 15 minutes

CODE REVIEW

- Mhàs
 - Learn
 - Share
 - Make code better

CODE REVIEW EXERCISE

What?

- Is it clear what the code is doing?
- Would it be easy to reuse parts of the program? Could this fail?
- Is it easy to test?
- Are there any bugs?
- Are function/variable names descriptive?

Hows

- Be positive and constructive
- Point out things you like
- Clarify reasoning where code doesn't match your expectations
- Discuss your opinions
- Is it worth changing? Sometimes it's OK to comment but take no action