

Things* I wish I knew* before starting a real job

TODO:explain *

Trevor Sullivan

February 23, 2018

Background in 30 seconds, GO!

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS
- Masters program? hell yeah!

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS
- Masters program? hell yeah!
- Internship with ayfie on fuzzy prefix search

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS
- Masters program? hell yeah!
- Internship with ayfie on fuzzy prefix search
- Job?

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS
- Masters program? hell yeah!
- Internship with ayfie on fuzzy prefix search
- Job?
- yes

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS
- Masters program? hell yeah!
- Internship with ayfie on fuzzy prefix search
- Job?
- yes
- search for lawyers

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS
- Masters program? hell yeah!
- Internship with ayfie on fuzzy prefix search
- Job?
- yes
- search for lawyers

Background in 30 seconds, GO!

- Came to UofAZ as a music/linguistics student
- Realized music is a dumb thing to spend \$40,000 on
- Hey computers are cool
- Double major in linguistics/CS
- Masters program? hell yeah!
- Internship with ayfie on fuzzy prefix search
- Job?
- yes
- search for lawyers

TIME

Overview

Overview

- I worked as an intern for ayfie for my hlt project on fuzzy prefix search (cf October 2016 Data Science Meetup)

Overview

- I worked as an intern for ayfie for my hlt project on fuzzy prefix search (cf October 2016 Data Science Meetup)
- I was hired afterwards as a Professional Service Engineer, to start in May

Overview

- I worked as an intern for ayfie for my hlt project on fuzzy prefix search (cf October 2016 Data Science Meetup)
- I was hired afterwards as a Professional Service Engineer, to start in May
- I took some notes on things* that I found particularly frustrating to not know* ahead of time

Overview

- I worked as an intern for ayfie for my hlt project on fuzzy prefix search (cf October 2016 Data Science Meetup)
- I was hired afterwards as a Professional Service Engineer, to start in May
- I took some notes on things* that I found particularly frustrating to not know* ahead of time

Overview

- I worked as an intern for ayfie for my hlt project on fuzzy prefix search (cf October 2016 Data Science Meetup)
- I was hired afterwards as a Professional Service Engineer, to start in May
- I took some notes on things* that I found particularly frustrating to not know* ahead of time

Starting any new job comes with a learning phase that lasts weeks or months, but knowing* about some things* ahead of time will help you make the transition more easily

On things and knowing

- Things

On things and knowing

- Things
 - This is just an excuse for me to bring up **impostor syndrome** because it didn't fit anywhere else in the talk

On things and knowing

- Things
 - This is just an excuse for me to bring up **impostor syndrome** because it didn't fit anywhere else in the talk
 - it never goes away

On things and knowing

- Things
 - This is just an excuse for me to bring up **impostor syndrome** because it didn't fit anywhere else in the talk
 - it never goes away
 - **NOW** is the time to practice your confidence and self esteem

On things and knowing

- Things
 - This is just an excuse for me to bring up **impostor syndrome** because it didn't fit anywhere else in the talk
 - it never goes away
 - **NOW** is the time to practice your confidence and self esteem
- Learning

On things and knowing

- Things
 - This is just an excuse for me to bring up **impostor syndrome** because it didn't fit anywhere else in the talk
 - it never goes away
 - **NOW** is the time to practice your confidence and self esteem
- Learning
 - learning *about*

On things and knowing

- Things
 - This is just an excuse for me to bring up **impostor syndrome** because it didn't fit anywhere else in the talk
 - it never goes away
 - **NOW** is the time to practice your confidence and self esteem
- Learning
 - learning *about*
 - *learning* learning

Table of contents

1. System Administration | The Command Line
2. Team workflows | Personal task management
3. Docker, containerization, packaging | Git/Git-flow
4. Continuous Integration | Testing
5. Cluster computing | Functional programming

System Administration | The Command Line

Be aware of: Finer points of sysadmin

Be aware of: Finer points of sysadmin

When you're working on a server, you will need to learn some particular skills and methods to manage things

Managing volumes `df -h`, `fdisk`, `mkfs`

Be aware of: Finer points of sysadmin

When you're working on a server. you will need to learn some particular skills and methods to manage things

Managing volumes `df -h`, `fdisk`, `mkfs`

ssh tunnels and UDP `ssh -L 18080:127.0.0.1:18085 -R 18086:127.0.0.1:18086 user@server`

Be aware of: Finer points of sysadmin

When you're working on a server. you will need to learn some particular skills and methods to manage things

Managing volumes `df -h`, `fdisk`, `mkfs`

ssh tunnels and UDP `ssh -L 18080:127.0.0.1:18085 -R
18086:127.0.0.1:18086 user@server`

web access control `nginx`

Do it now: The Command Line

- Navigating a *nix filesystem

Do it now: The Command Line

- Navigating a *nix filesystem
- text editor (vim, emacs, nano)

Do it now: The Command Line

- Navigating a *nix filesystem
- text editor (vim, emacs, nano)
- tmux or screen

They may seem dumb at first, but you will soon find you can't live without it

Like browsing without tabs

```
File Edit View Search Terminal Help
"clustersPerLevel": 20,
"documentsPerCluster": 1000,
"gramianSingularVectors": 50,
"gramianThreshold": 0.14,
"maxDfFraction": 0.25,
"maxRecursionDepth": 2,
"minDf": 3
    }
  },
  "type": "CLUSTERING"
},
"details": {
  "applicationId": "app-20180215165851-0005",
  "description": "Execution completed successfully",
  "timestamp": "2018-02-15T17:05:54.0368Z"
},
"id": "1727374435",
"state": "SUCCEEDED"
}
]
},
"_links": {
  "self": {
    "href": "http://localhost:93/ayfie/v1/jobs"
  }
}
}

dev@docker2:/home/docker-demos/compositions/ayfie/new-wiki-pedia$
[trevor] 0:ssh 1:bahr-monitor 2:bash* 3:bash- "docker2" 23:19 15-Feb-18
```

Team workflows | Personal task management

Be aware of: Scrum and Team Workflows

- Every company does this a different way

Be aware of: Scrum and Team Workflows

- Every company does this a different way
- scrum master

Be aware of: Scrum and Team Workflows

- Every company does this a different way
- scrum master
- sprint review

Be aware of: Scrum and Team Workflows

- Every company does this a different way
- scrum master
- sprint review
- board

Boards

Professional Services

Kanban board

Release ▾ ***

Q Quick filters ▾ Trevor Sullivan ▾

TO DO 2 of 23 BLOCKED 5 of 12 IN PROGRESS 2 of 9 DONE 4 of 6

▼ Bahr POC 5 issues

Enable deeplink functionality

Bahr POC

🟢 ⬆️ ⬆️ ⬆️ ⬆️ PSE-144 👤

bahr run bigger collection with new extractors

Bahr POC

🟢 ⬆️ ➡️ PSE-196 👤

classify new bahr, but with pathclassifications

Bahr POC

🟢 ⬆️ ➡️ PSE-198 👤

cluster new bahr

Bahr POC

🟢 ⬆️ ➡️ PSE-197 👤

▼ BryanCave POC 4 issues

▼ Internal Stuff 1 issue

▼ Issues without epics 3 issues

read suggest source

🟢 ⬆️ ➡️ PSE-195 👤

wikipedia classification and clusering optimization

🟢 ⬆️ ⬆️ ⬆️ ⬆️ PSE-95 👤

PSE-174 Write tests for axelcing scripts

regression tests for logical functions in feeday/le.py

🟢 ⬆️ ➡️ PSE-175 👤

Boards

CRUNCH

AP Sprint 65

Stabilize for iConnect

🕒 3 days remaining Complete sprint ⋮

Quick Filters

Assignee

TO DO

IN PROGRESS

DONE

✓ AP-3034 DONE 2 sub-tasks Basic Spark Telemetry

Fix it

None

🔍 ↑

AP-3130

Changelog

None

🔍 ↑

AP-3137

✗ AP-2493 TO DO 2 sub-tasks Find out why some emails are missing thread meta.

Fix it in ultrakit-frontend

None

🔍 ↑

AP-3131

Investigate

None

🔍 ↑

AP-3163

Do it now: Task Management

- People tend to overestimate their own ability to remember things

Do it now: Task Management

- People tend to overestimate their own ability to remember things
- There is a limited amount of space in your working memory, and to commit something to long term memory, you need to call it back to working memory several times

Do it now: Task Management

- People tend to overestimate their own ability to remember things
- There is a limited amount of space in your working memory, and to commit something to long term memory, you need to call it back to working memory several times
- This is a waste of brain power. Everything you have to remember distracts you.

Do it now: Task Management

- People tend to overestimate their own ability to remember things
- There is a limited amount of space in your working memory, and to commit something to long term memory, you need to call it back to working memory several times
- This is a waste of brain power. Everything you have to remember distracts you.
- Strive therefore to get things out of your head as soon as possible

Apps: Asana

The screenshot displays the Asana web application interface. On the left is a dark sidebar with navigation links: 'My Favorites' (with sub-links: ingbot, Elymachine, blog, move, syfe), 'Reports', 'PROJECTS' (with sub-links: Incubation, ingbot, blog, Elymachine, move, syfe, 352:UNIX & C, APL, CPR, CSC Computer Organization, CompLing, CompLing), 'Episode 1 finalized', 'Episode 2 done', 'Episode 3 done', 'Elymachine', 'German', 'History', 'Laptop purchased', 'Mimemini Machine', 'Moon and Planets', 'Persian', and 'Podcast'. The main content area is titled 'My Tasks' and includes a search bar and tabs for 'List', 'Calendar', and 'Files'. The 'List' tab is active, showing a task list. The tasks are organized into sections: 'Today' and 'Upcoming'. Each task has a checkbox, a description, and a due date. The 'Today' section includes tasks like 'setup up application.yml that does email threading and reduplication and measure performance' and 'Look at Dene yenisian evidence'. The 'Upcoming' section includes tasks like 'setup two amazon instances' and 'ingest and process error dataset, measure time'. A task detail view is open on the right, showing the task 'Look at Dene yenisian evidence' with a description, a list of assignees (Trevor Sullivan), and a comment section.

My Tasks **My Personal Tasks**

Today

- ✓ setup up application.yml that does email threading and reduplication and measure performance
- ✓ set up same app on a server and minimize resource requirements
- ✓ Look at [Dene yenisian](#) evidence
- ✓ pay off credit card
- ✓ pay cox
- ✓ pay tlp
- ✓ contacts;

Upcoming

- ✓ setup two amazon instances
- ✓ ingest and process error dataset, measure time
- ✓ do several runs to stabilize numbers
- ✓ find maximum amount to data ingestible using apache email dataset
- ✓ setup a demo based on [wikipedia](#) content
- ✓ fill out check writing thing for cash management account
- ✓ Look into active learning for [Elymachine](#)
- ✓ [Xbox](#) smelly parenthetical
- ✓ Make new survey data into .ped format
- ✓ make [gu's](#) changes
- ✓ data normalization: why it's important to have input values in a known range
- ✓ respond to this
- ✓ make twitch streaming [gettings](#) in obs
- ✓ make twitch
- ✓ [hackyhouse](#) share photo
- ✓ Understanding LSTM Networks -- colah's blog
- ✓ [rip read] Pwd: Augur: Mining Human Behaviors from Fiction to Power Interactive Systems
- ✓ skinner box games
- ✓ The operation trained for the worst world

Task Detail: Look at Dene yenisian evidence

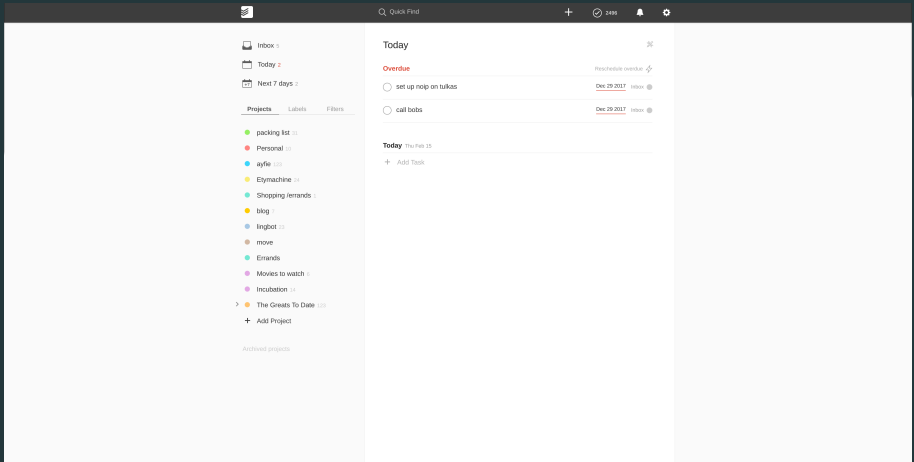
Description

Trevor Sullivan created task - May 28, 2017
Trevor Sullivan assigned to you - May 28, 2017
Trevor Sullivan marked today - May 28, 2017

Write a comment...

Followers

Apps: Todoist



Apps: Trello

Hannibal/Scipio2



The Greats to Date

Free



Team Visible

Hannibal/Scipio II

Put both files on same machine

Line up and do meticulous primary edits

listen and make notes for secondary edits

listen at super speed and make notes for tertiary edits

Upload to media host

Update Blog

publish to RSS

Add a card...

in progress

Add a card...

done

Read relevant book

🕒 Jul 11, 2016

Make notes for show

🕒 Jul 12, 2016 🗨 1

Do show

🕒 Jul 21, 2016

Do show

Add a card...

Add a list...

Apps: todo.txt

```
todo.todo x
1
2
3 ▼Stuff to do now:
4   ◦ do CV test with syllables
5   ◦ add language names to stopwords list
6   ◦ add code to submission
7   ◦ add readme to submission (python notebook)
8
9   ◦ run through naive bayes
10  ◦ run through k-nearest neighbors
11  ◦ add some neural network stuff to this list ◦ run through naive bayes
12  ◦ run through k-nearest neighbors
13  ◦ add some neural network stuff to this list ◦ run through naive bayes
14  ◦ run through k-nearest neighbors
15  ◦ add some neural network stuff to this list
16
17
18 ▼features:
19   ✓ bag of words @done (16-11-21 13:25)
20   ✓ characters present @done (16-11-23 12:30)
21   ◦ syllables
22   ◦ bigrams (over characters and/or syllables)
23   ✓ attested year (or none) @done (16-11-28 16:49)
24   ◦ normalize over centuries, or continuous value
25   ◦
26
27 ▼Machine learning categorization:
28   ✓ vectorize definitions (features below) @done (16-11-23 12:30)
29   ✓ SVM @today @done (16-11-28 14:49)
30   ✓ test different kernels @done (16-11-28 19:44)
31   ✓ Make a confusion matrix visualization @done (16-11-29 08:46)
32
33   ◦ run through naive bayes
34   ◦ run through k-nearest neighbors
35   ◦ add some neural network stuff to this list
36
37
38 ▼After project is done:
```

Apps: Bulletjournal

15th

- ▲ pick up library book on hold
- Amy for lunch @ 11:30 am
- ☐▼ make summer bucket list
- started watching Vikings finally

16th

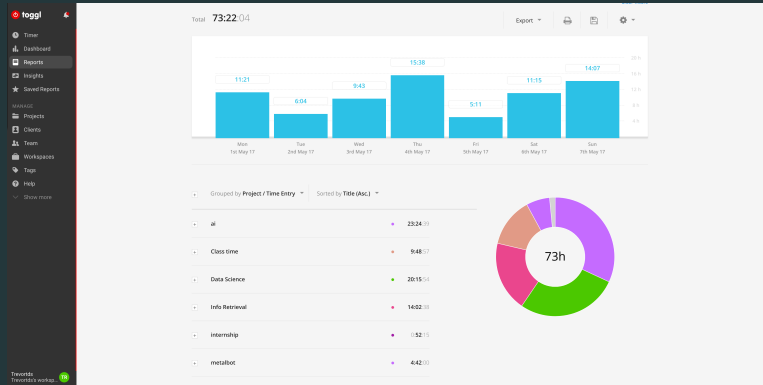
- look for UNC tickets

17th

- ▲ OT @ 3pm
- ▲ dinner at Mellow w/ Carters @ 5:30pm
- finish Friday's blog post
- Sam said, "Mom, I can't be quiet, I'm a talking"

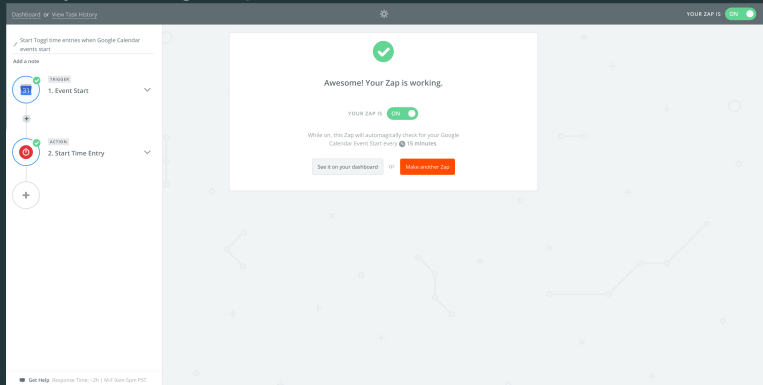
Advanced: Time tracking

Not essential, but if you already have a good system for keeping track of your tasks.



Super-Advanced: Automation

Many of the aforementioned apps have APIs that can be manipulated manually, or through Zapier



**Docker, containerization,
packaging | Git/Git-flow**

Be aware of: Docker, containerization, packaging

Maybe consider: `yaml` for configuration

Do it now: Git - not just for collaboration!

Continuous Integration | Testing

Be aware of: Continuous Integration

Do it now: Testing

Cluster computing | Functional programming

Be aware of: Cluster computing

Do it now: Functional programming