TEACHING SOFTWARE ENGINEERING FOR AI-ENABLED SYSTEMS

Christian Kaestner and Eunsuk Kang

Carnegie Mellon University

https://ckaestne.github.io/seai/

ICSE-SEET 2020

SOFTWARE ENGINEERING FOR ML-ENABLED SYSTEMS

Building, operating, and maintaining software systems with machine-learned components

SE 4 ML-SYSTEMS != TRAINING MODELS

G4 playground.ipynb

File Edit View Insert Runtime Tools Help Last edited on April 4

Comment Share

+ Code + Text Connect

1096	4	12	26	3	2	0
235	4	4	23	1	2	0

525 rows × 6 columns

```
# learning a classifier whether the result will be nonZero

from sklearn import tree

classifier=tree.DecisionTreeClassifier(max_depth=8)
classifier=classifier.fit(Xtrain,ynztrain)

print(classifier.score(Xtrain, ynztrain))
print(classifier.score(Xtest, ynztest))
```

```
0.8266666666666667
0.7295238095238096
```

```
# learning a regression model only on the nonZero data (test is on all data and somewhat

from sklearn import tree
```

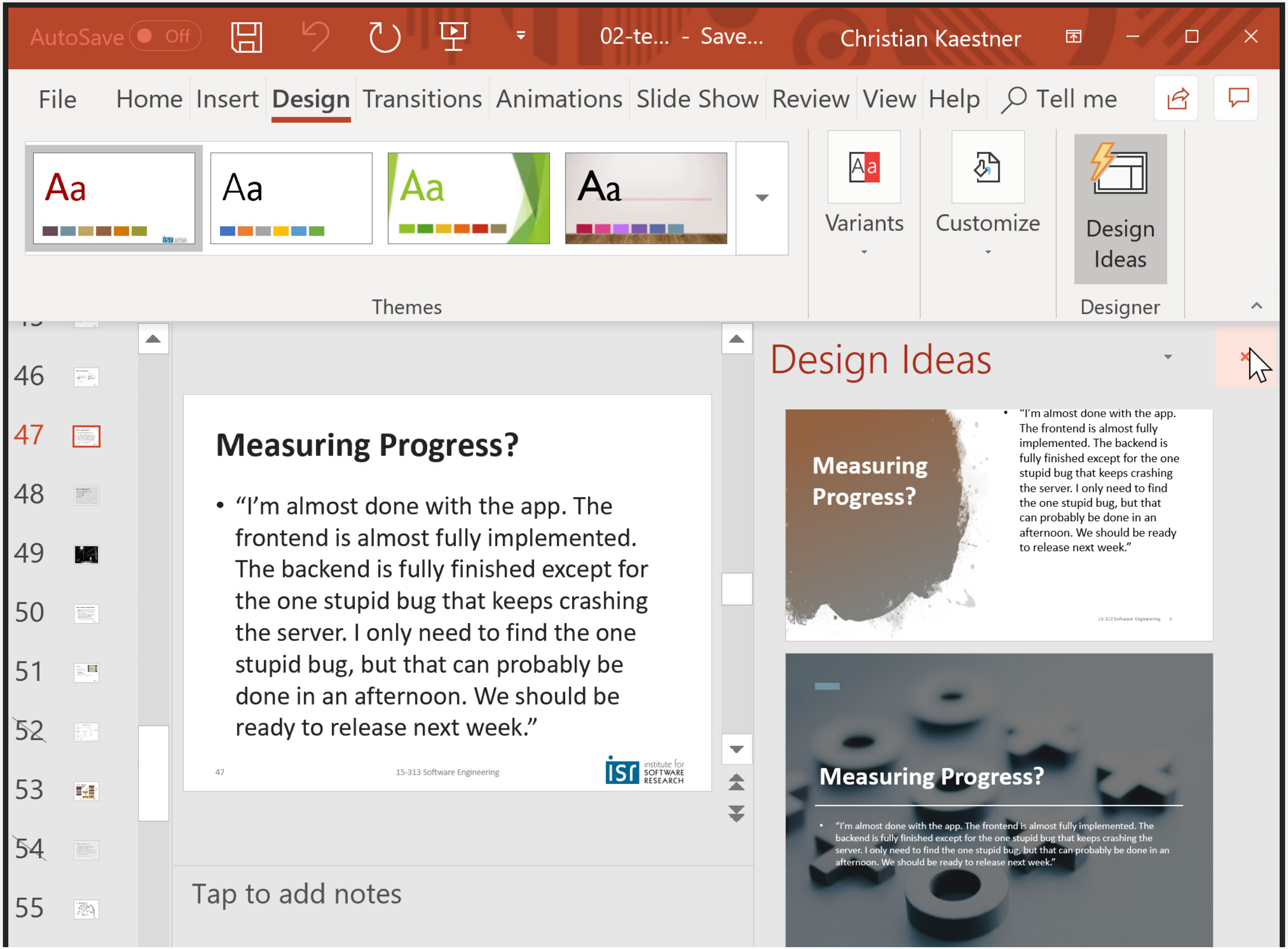
```python
predictor=tree.DecisionTreeRegressor(max_depth=8)
predictor=predictor.fit(XnzTrain,YnzTrain)


print(predictor.score(XnzTrain, YnzTrain))
print(predictor.score(Xtest, ytest))
```

```
0.9376379365613154
-2.437397740412892
```

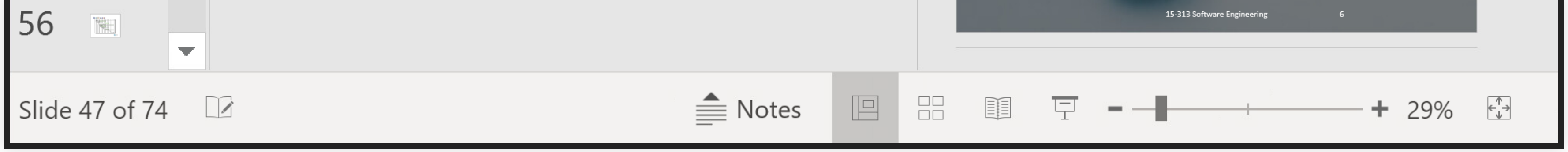
SE 4 ML-SYSTEMS != ML 4 SE

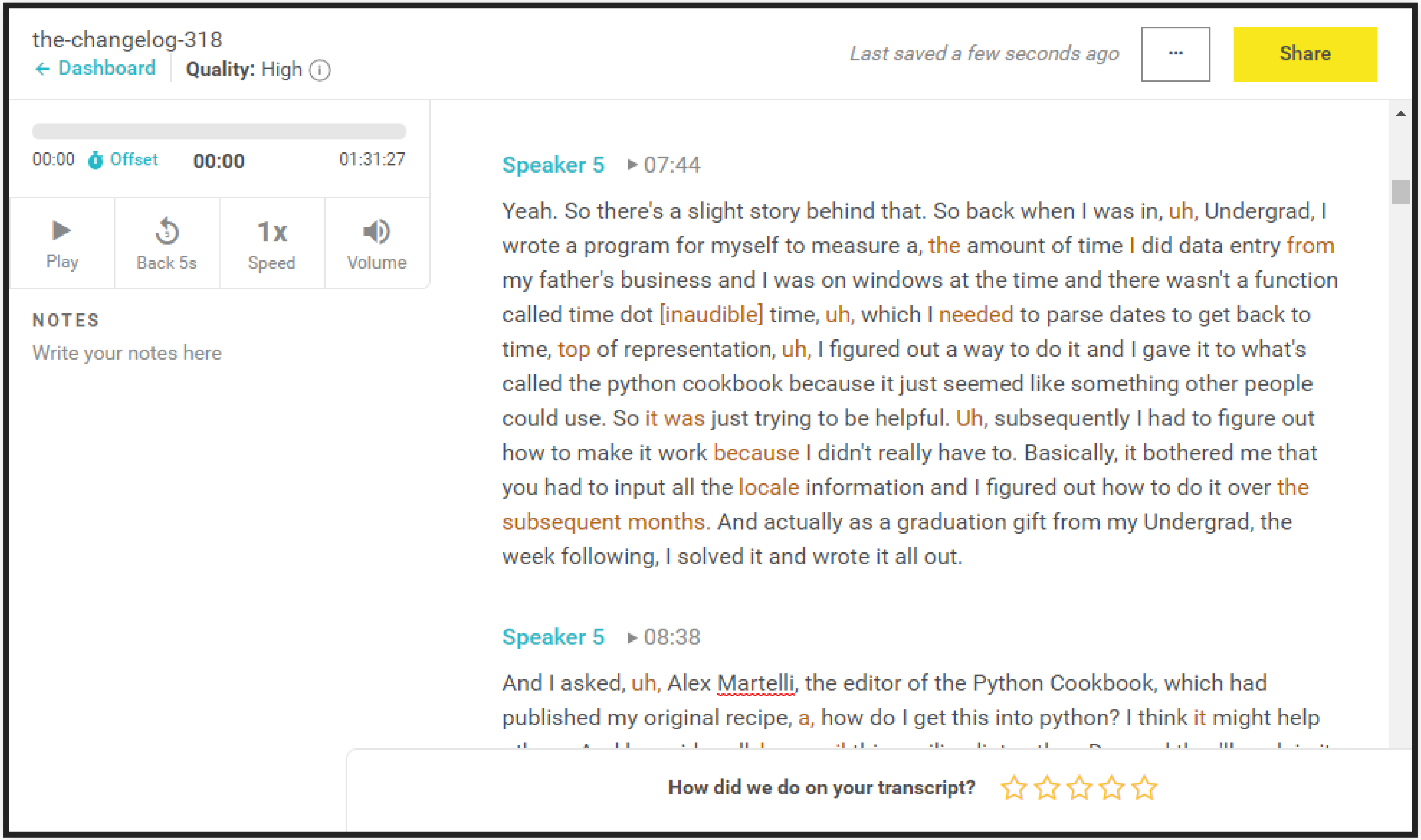
```python
import numpy as np

start = -1
stop = 1

x = np.lins
```

f	linspace	function
f	linspace(start, stop)	function
f	linspace(stop, start)	function
f	linspace(start, stop, sto…	function

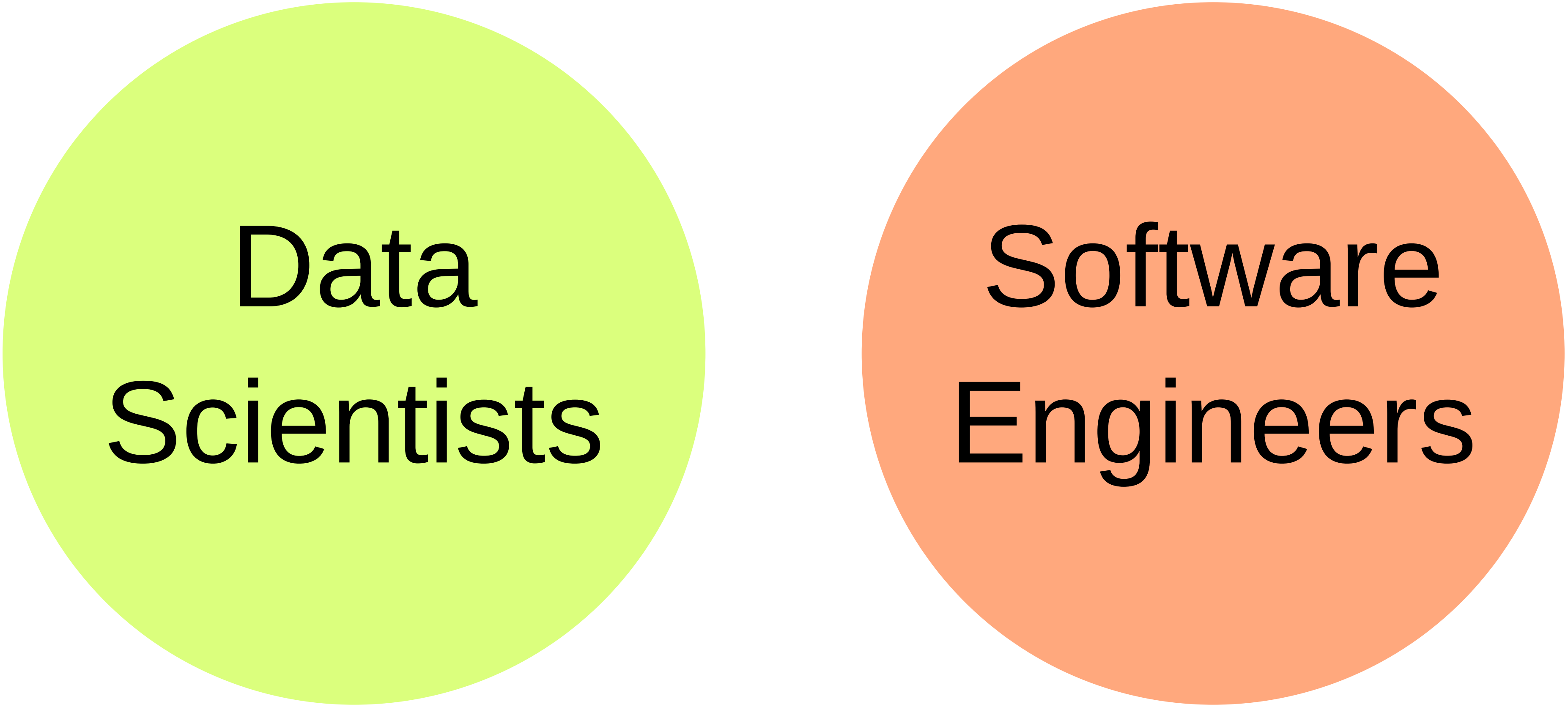


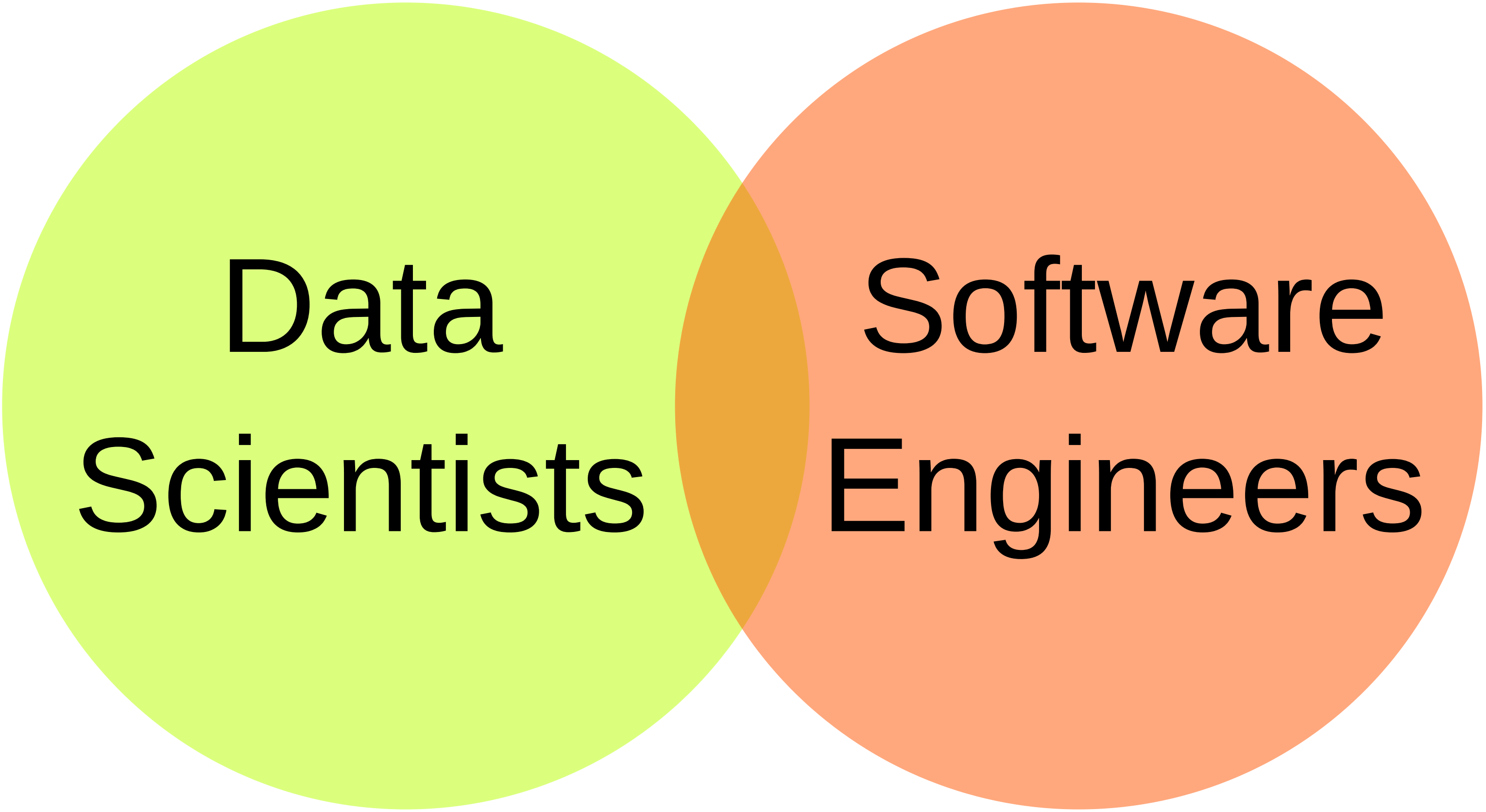
AutoSave Off
02-te... - Save...
Christian Kaestner
File
Home
Insert
Design
Transitions
Animations
Slide Show
Review
View
Help
Tell me
Themes
Variants
Customize
Design Ideas
Designer
Measuring Progress?
• "I'm almost done with the app. The frontend is almost fully implemented. The backend is fully finished except for the one stupid bug that keeps crashing the server. I only need to find the one stupid bug, but that can probably be done in an afternoon. We should be ready to release next week."
47
15-313 Software Engineering
isr institute for SOFTWARE RESEARCH
Tap to add notes
Design Ideas
Measuring Progress?
• "I'm almost done with the app. The frontend is almost fully implemented. The backend is fully finished except for the one stupid bug that keeps crashing the server. I only need to find the one stupid bug, but that can probably be done in an afternoon. We should be ready to release next week."
Measuring Progress?
• "I'm almost done with the app. The frontend is almost fully implemented. The backend is fully finished except for the one stupid bug that keeps crashing the server. I only need to find the one stupid bug, but that can probably be done in an afternoon. We should be ready to release next week."
46
47
48
49
50
51
52
53
54
55


56
15-313 Software Engineering
6
Slide 47 of 74
Notes
29%



the-changelog-318
Dashboard
Quality: High
Last saved a few seconds ago
...
Share
00:00
Offset
00:00
01:31:27
Play
Back 5s
1x
Speed
Volume
NOTES
Write your notes here
Speaker 5 07:44
Yeah. So there's a slight story behind that. So back when I was in, uh, Undergrad, I wrote a program for myself to measure a, the amount of time I did data entry from my father's business and I was on windows at the time and there wasn't a function called time dot [inaudible] time, uh, which I needed to parse dates to get back to time, top of representation, uh, I figured out a way to do it and I gave it to what's called the python cookbook because it just seemed like something other people could use. So it was just trying to be helpful. Uh, subsequently I had to figure out how to make it work because I didn't really have to. Basically, it bothered me that you had to input all the locale information and I figured out how to do it over the subsequent months. And actually as a graduation gift from my Undergrad, the week following, I solved it and wrote it all out.
Speaker 5 08:38
And I asked, uh, Alex Martelli, the editor of the Python Cookbook, which had published my original recipe, a, how do I get this into python? I think it might help
How did we do on your transcript?


Data Scientists
Software Engineers

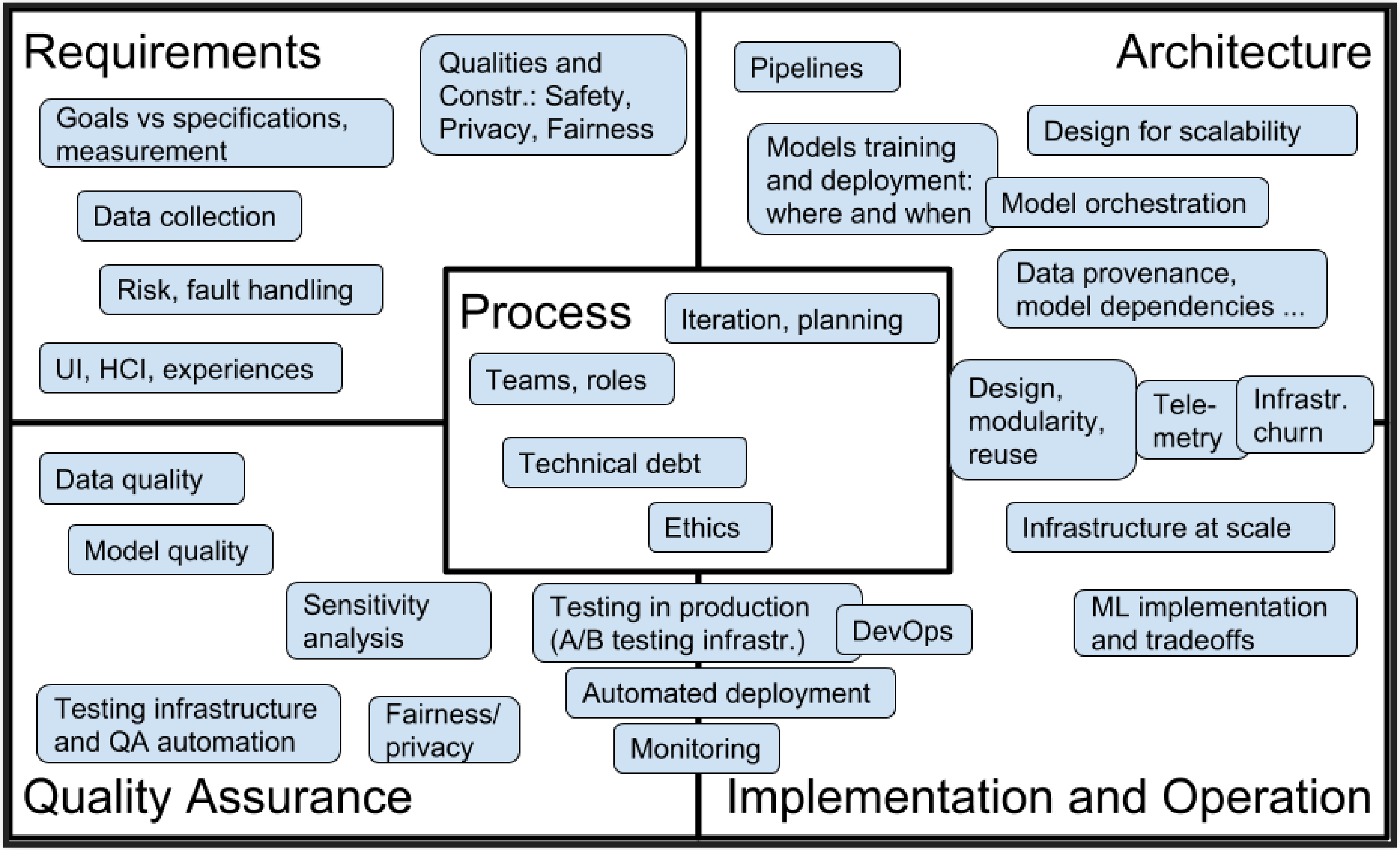

Data Scientists
Software Engineers

MANY CHALLENGES

- Missing specifications
- Environment is important (feedback loops, data drift)
- Nonlocal and nonmonotonic effects
- Testing in production
- Data management, versioning, and provenance
- Fairness, robustness, interpretability

SOFTWARE ENGINEERS CAN CONTRIBUTE

- Missing specifications -- *implicit, vague specs very common; safe systems from unreliable components, risk analysis* ("ML is requirements engineering")
- Environment is important -- *the world vs the machine* (paper)
- Nonlocal and nonmonotonic effects -- *feature interactions, system testing*
- Testing in production -- *continuous deployment, A/B testing*
- Data management, versioning, and provenance -- *stream processing, event sourcing, data modeling*
- Fairness, robustness, interpretability -- *traditional requirements engineering questions*



https://github.com/ckaestne/seai

OUR VIEW

While developers of simple traditional systems may get away with poor practices, most developers of AI-enabled systems will not.

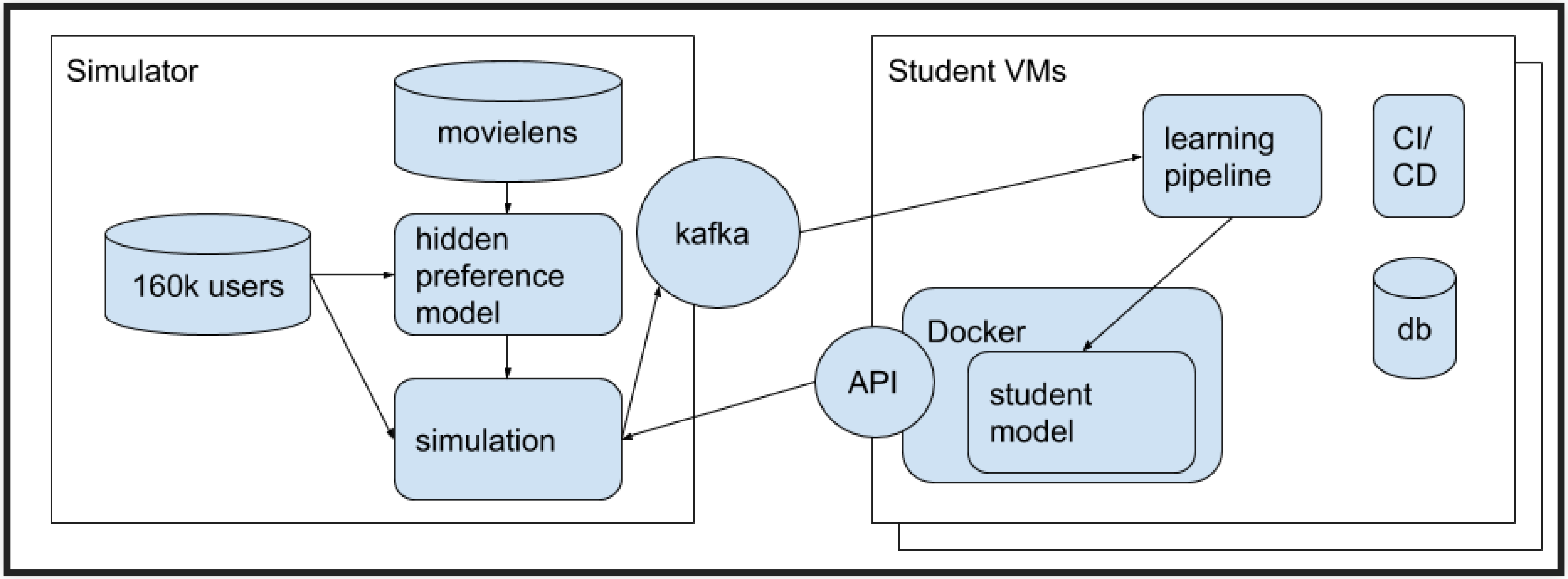
This is an education problem, more than a research problem.

ASSIGNMENTS

Break the habit of modeling in notebooks on static datasets

Design for realistic "production" setting: deployment, experimentation in production, data drift and feedback loops

Movie recommendations for 1 million simulated users in real time



SUMMARY: SOFTWARE ENGINEERING FOR ML-ENABLED SYSTEMS

- Building, operating, and maintaining systems with ML component
- Data scientists and software engineers have different expertise, both needed
- Software engineering view on intelligent *systems*:
    - User interaction design
    - Model qualities and deployment tradeoffs
    - Risk analysis and safety
    - Architecture, deployment, telemetry design
    - Quality assurance, fairness, robustness, interpretability
- All material, including slides, assignments, video recordings, shared under Creative Commons license

https://ckaestne.github.io/seai/