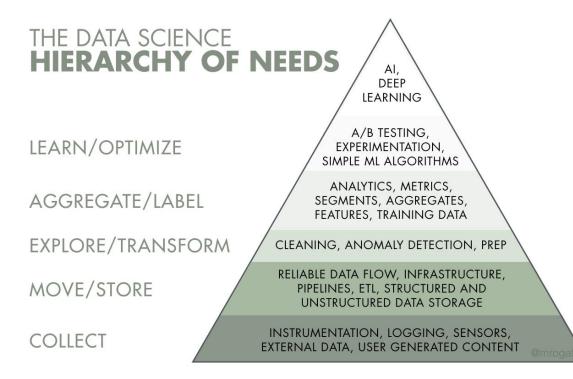
Innovating on software development

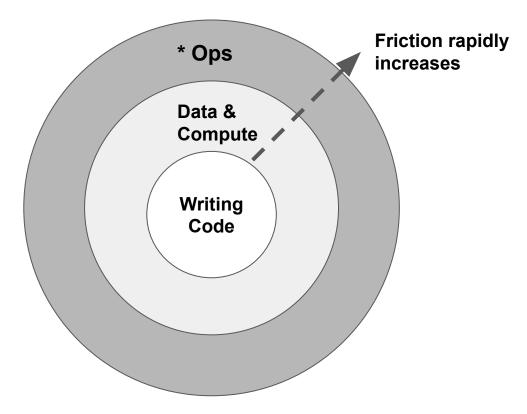
Some observations about the current state of software development

We need to write code at all stages

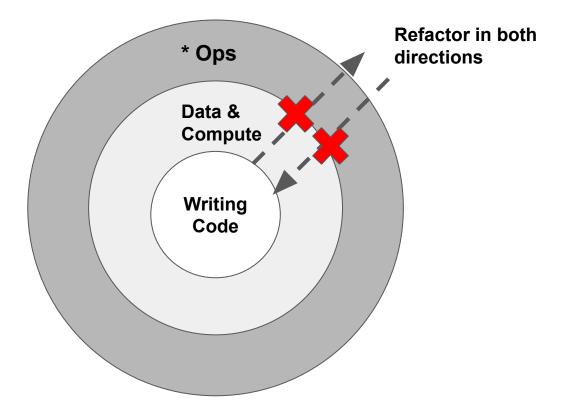


Monica Rogati's Data Science Hierachy of Needs

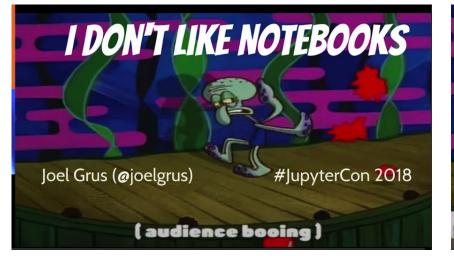
Innovation on tools are focused on the outer loops

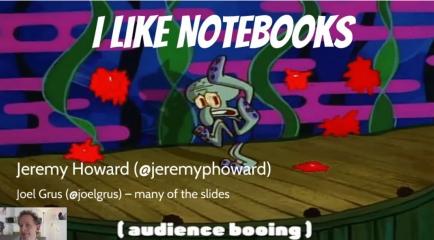


We Live in Two Worlds



How We Write Code is Hotly Debated





How We Write Code is Hotly Debated

Operationalizing Machine Learning: An Interview Study

Shreya Shankar*, Rolando Garcia*, Joseph M. Hellerstein, Aditya G. Parameswaran University of California, Berkeley {shreyashankar,rogarcia,hellerstein,adityagp}@berkeley.edu

*Co-first authors

······

Strong Opinions on Jupyter Notebooks. Participants described strongly opinionated and different philosophies with respect to how to use Jupyter notebooks in their workflows. Jupyter notebooks were heavily used in development to support high velocity, which we did not find surprising. However, we were surprised that although participants generally acknowledged worse code quality in notebooks, some participants preferred to use them in production to minimize the differences between their development and production environments. P6 mentioned that they could debug quickly when locally downloading, executing, and manipulating data from a production notebook run. P18 remarked on the modularization benefits of a migration from a single codebase of scripts to notebooks:

We put each component of the pipeline in a notebook, which has made my life so much easier. Now [when debugging], I can run only one specific component if I want, not the entire pipeline... I don't need to focus on all those other components, and this has also helped with iteration. On the other hand, some participants strongly disliked the idea of notebooks in production (P10, P15). P15 even went as far as to philosophically discourage the use of notebooks in the development environment: "Nobody uses notebooks. Instead, we all work in a shared code base, which is both the training and serving code base and people kick off jobs in the cloud to train models." Similarly, P10 recounted a shift at their company to move any work they wanted to reproduce or deploy out of notebooks:

How We Write Code is Hotly Debated

Operationalizing Machine Learning: An Interview Study

Shreya Shankar*, Rolando Garcia*, Joseph M. Hellerstein, Aditya G. Parameswaran University of California, Berkeley {shreyashankar,rogarcia,hellerstein,adityagp}@berkeley.edu *Co-first authors

The anecdotes on notebooks identified conflicts between competing priorities: (1) Notebooks support high velocity and therefore need to be in development environments, (2) Similar development and production environments prevents new bugs from being introduced, and (3) It's easy to make mistakes with notebooks in production, e.g., running with the wrong inputs; copy-pasting instead of reusing code. Each organization had different rankings of these priorities, ultimately indicating whether or not they used notebooks in production.

Two Axes Of Innovation

Literate Programming

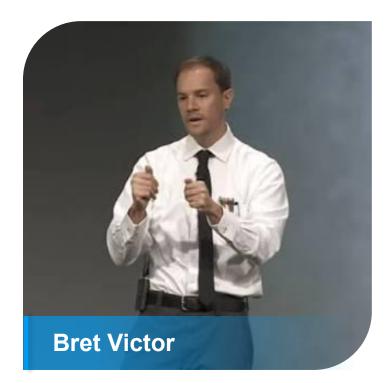
Combines programming and documentation. Software is written to be read by humans. Natural language interspersed with code.

Exploratory Programming

Having an immediate connection and feedback to your code. Interactive, dynamic development

Prior Art

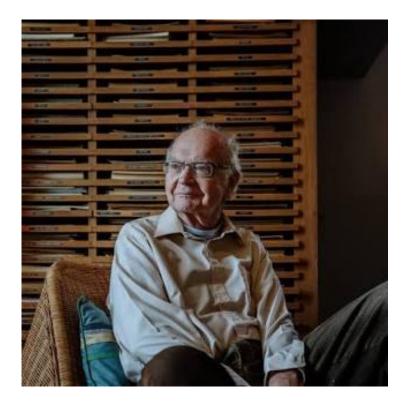




Literate Programming

- Combines programming and documentation.
- Software is written to be read by humans.
- Natural language interspersed with code.

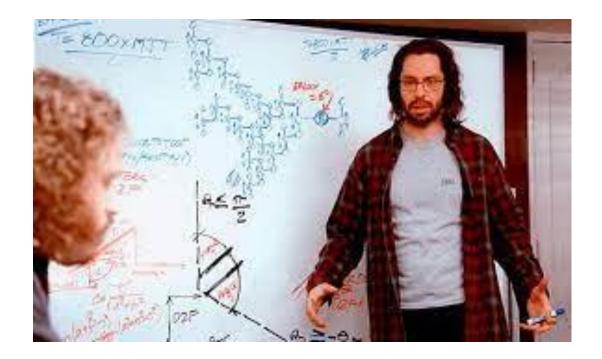
Knuth: "The Yoda Of Silicon Valley"





The basis for an engineering rite of passage



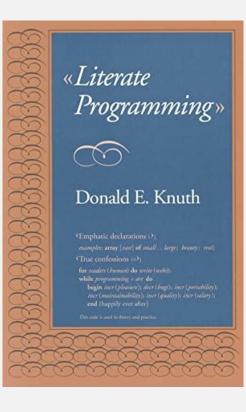


One Of The Most Celebrated Works In Computer Science

If you think you're a really good programmer . . . read [Knuth's] Art of Computer Programming. . . . You should definitely send me a résumé if you can read the whole thing. —Bill Gates

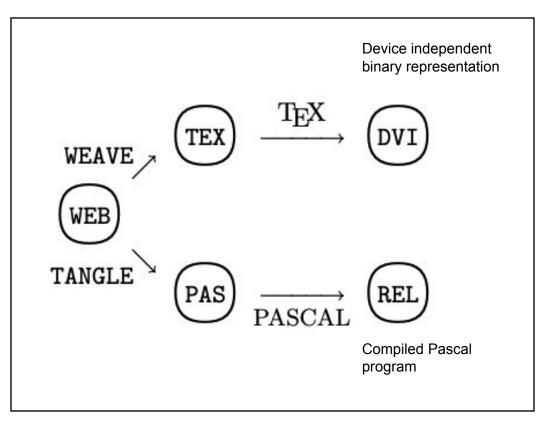
Programming" is the Bible of its field. "Like an actual bible, it is long and comprehensive; no other book is as comprehensive," said Peter Norvig, a director of research at Google. After 652 pages,

Lesser Known: Coined The Term "Literate Programming"

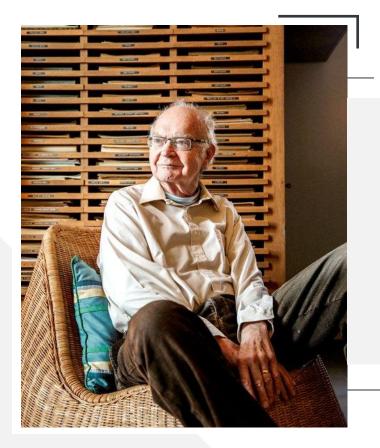


A collection of work and papers discussing Literate Programming.

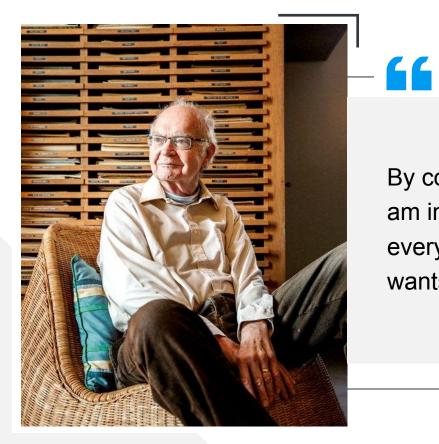
WEB: A Literate Programming Language



One file generates source code (Pascal) and documentation (Tex) from it

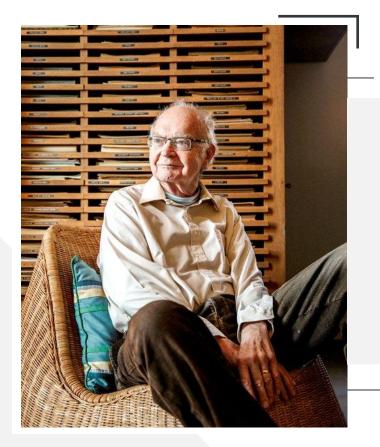


Let us change our traditional attitude to the construction of programs: Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on explaining to human beings what we want a computer to do.

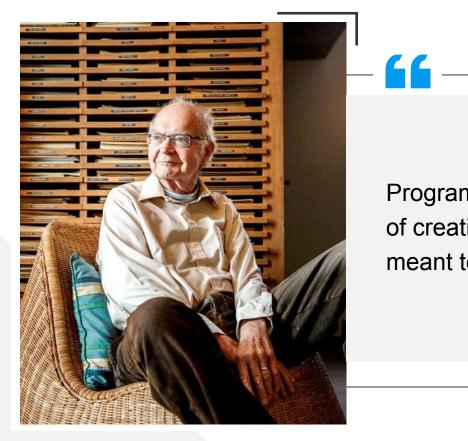


By coining the phrase "literate programming," I am imposing a moral commitment on everyone who hears the term; surely nobody wants to admit writing an illiterate program.

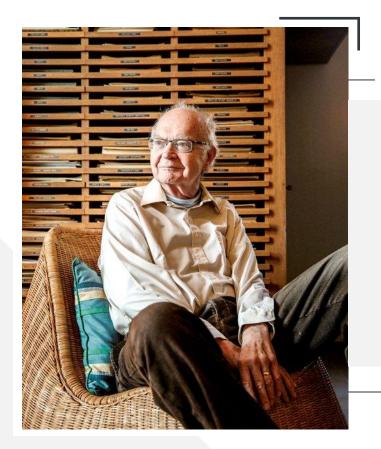
66 -



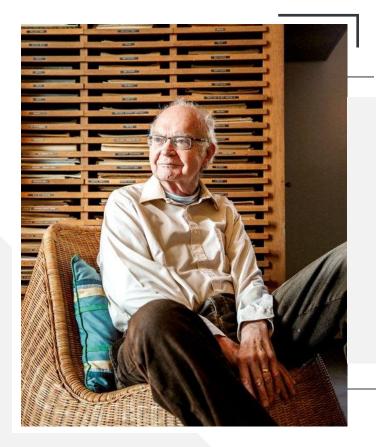
<u>All</u> of the major problems associated with computer programming ...are ameliorated when programs and their dialogs with users become more literate.



Programming is best regarded as the process of creating works of literature, which are meant to be read.

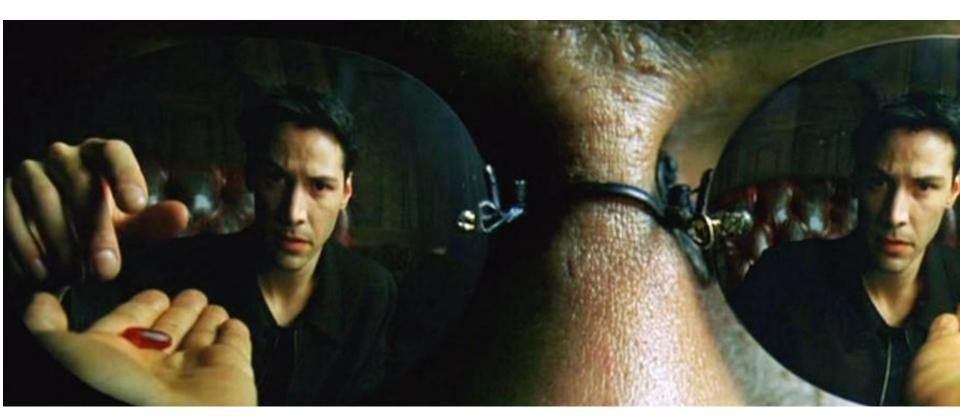


In retrospect, the fact that a "literate" program takes much less time to debug is not surprising, because the WEB language encourages a discipline that I was previously unwilling to impose on myself.

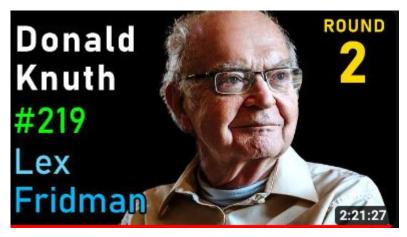


I have found to my surprise that the total time of writing and debugging a WEB program is no greater than the total time of writing and debugging an ALGOL or PASCAL program, even though my WEB programs are much better, and [have] substantially more documentation

Knuth: We Need To Change The Way We Think About Programming



He hasn't changed his mind

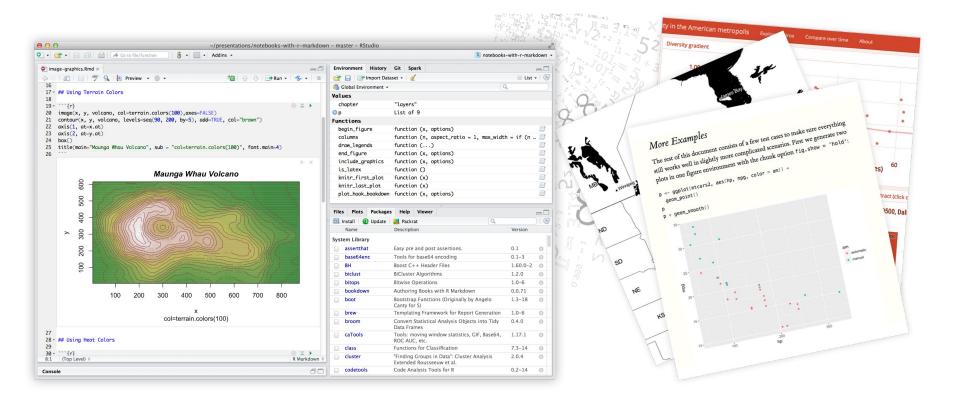


September 2021: Interview

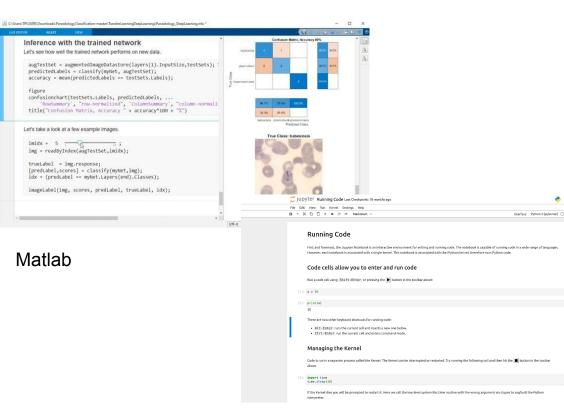


December 2022: Algorithms Lecture

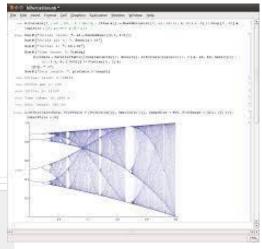
Nod To Literate Programming: R Markdown



Nod To Literate Programming: Notebooks



figure

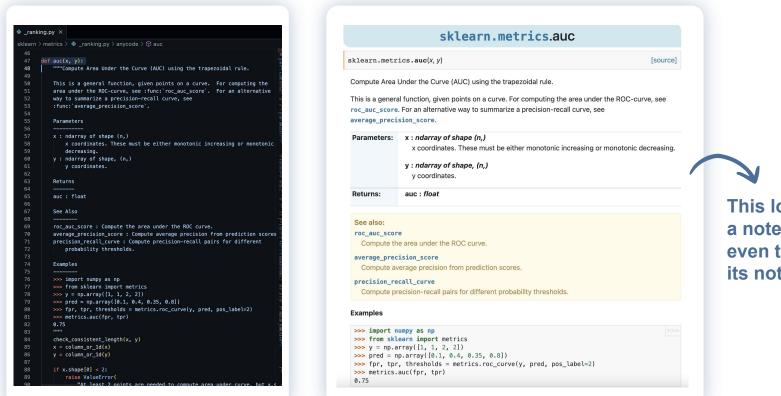


2

Mathematica

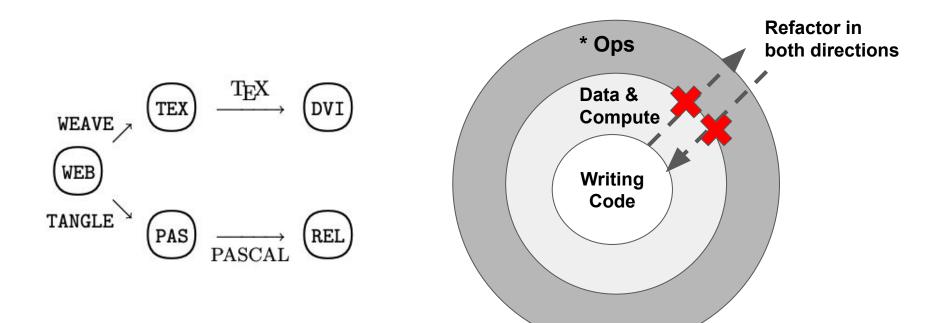
Jupyter

Writing prose + "production" code is painful



This looks like a notebook, even though its not.

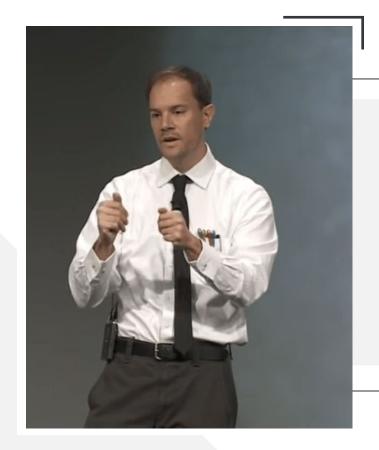
Literate Programming vs Notebooks: Still A Gap



Exploratory Programming

- Having an <u>immediate</u> connection and feedback to your code.
- Interactive, dynamic development

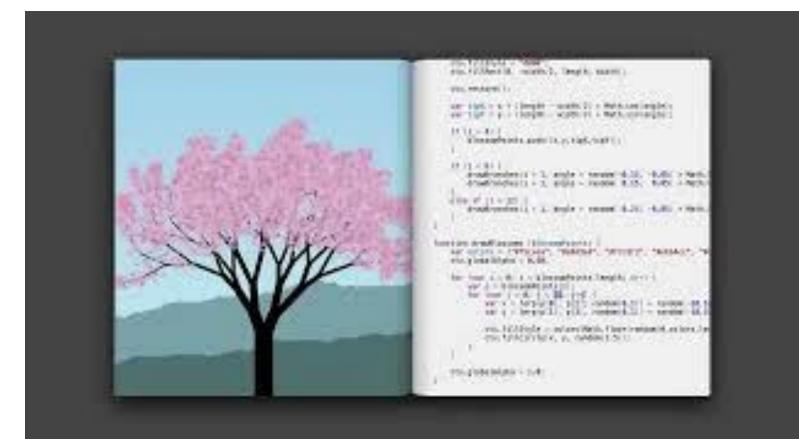
Bret Victor: Champion of Exploratory Programming



"

I wanna make a new MatLab or a cross between MatLab and Garageband.

Bret Victor: Inventing On Principle



What If We Combined Literate & Exploratory Programming?

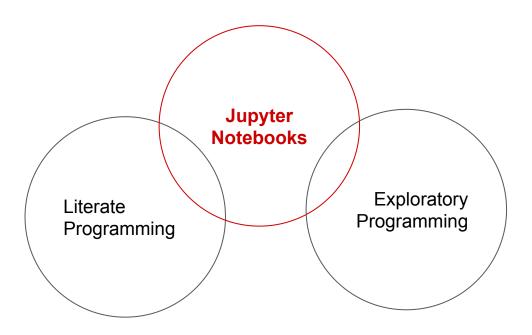
nbdev.fast.ai

Create delightful software with Jupyter Notebooks

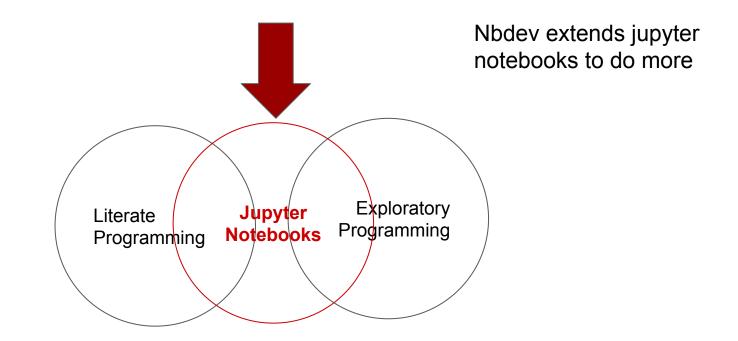
Write, test, document, and distribute software packages and technical articles — all in one place, your notebook.

Get started upyte GitHub + quarto Deck.pop pages-build-deployment Back non (indexn-) O build tos A Deck of cards is a collection of Card objects: Ortails deck = Deck(deck O deploy A&; 2&; 3&; A : 2 : 3 ; 5 \$\$ 6 \$\$ 7 \$\$ 8 \$\$ 9 \$\$ 10 \$\$ 35 0\$ 7\$ \$\$ 5 \$\$ 6 \$\$ 7 \$\$ 8 \$\$ 9 \$\$ 10 \$\$ 35 0\$ \$\$ 7\$ \$\$ 5 \$\$ 6 \$\$ 7\$ \$\$ 8 \$\$ 9\$ \$\$ 10 \$\$ 35 0\$ \$\$ \$\$ 5 \$\$ 6 \$\$ 7\$ \$\$ 8\$ 9\$ \$\$ 10 \$\$ 35 0\$ \$\$ \$\$ deck.pop There are 52 cards in a deck test_eq(len(deck), 52) 🕜 рурі ... cards-deck 0.0.3 fastai / packages / cards-deck 0.0.3 🖪 🛷 🤗 🚖 📀 deck.por 1.4 In example repository for ribdev Conta Files Labels Bedges There are 51 cards left in the deck now. License: Apache Software
 Home: https://fastol.github.ic test eg(len(deck), 51) Orweisperent Impschlesse gimelsio
 Documentation Interschlessi gimelsio
 7 total downiseds
 Lett upliced 14 hours and 56 minutes age Project description Deck of Cards

nbdev

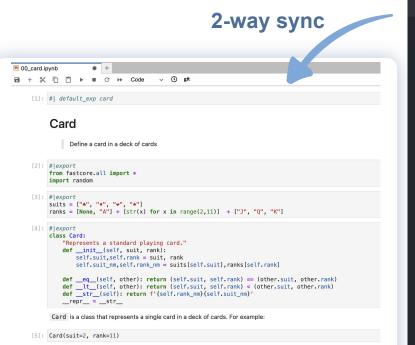


nbdev



What is nbdev?

1. Generate Python Modules



[5]: **J**♥

Creating Cards

We can create carde like this

🛑 🔵 🛑 🍦 card.py

AUTOGENERATED! DO NOT EDIT! File to edit: ../00_card.ipynb.

%% auto 0
__all__ = ['suits', 'ranks', 'Card']

%% ../00_card.ipynb 2
from fastcore.all import *
import random

‰ ../00_card.ipynb 3 suits = ["♠", "♦", "♥", "♠"] ranks = [None, "A"] + [str(x) for x in range(2,11)] + ["J", "Q", "K"]

%% ../00_card.ipynb 4
class Card:
 "Represents a standard playing card."
 def __init__(self, suit=0, rank=2):
 self.suit,self.rank = suit, rank
 self.suit_nm,self.rank_nm = suits[self.suit],ranks[self.rank]

def __eq__(self, other): return (self.suit, self.rank) == (other.suit, other.rank
def __lt__(self, other): return (self.suit, self.rank) < (other.suit, other.rank)
def __str__(self): return f'{self.rank_nm}{self.suit_nm}'
__repr__ = __str__</pre>

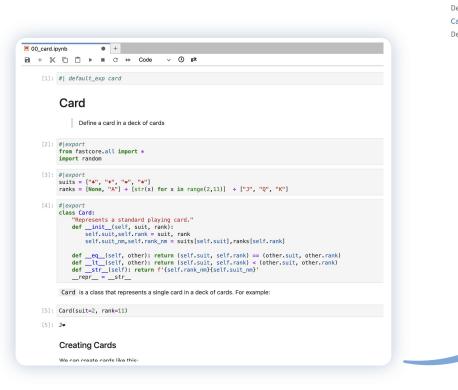


Use your favorite IDE for nav & edits

2. Develop Modules From Notebooks

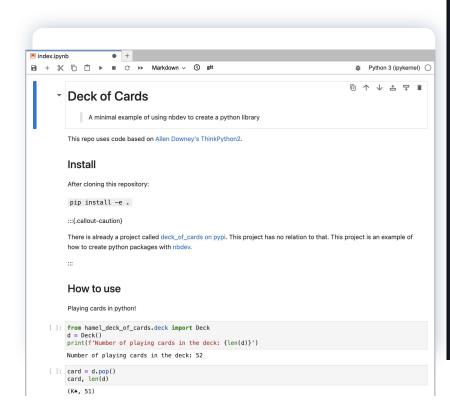
	O1_deck.ipynb × +
■ 00_card.ipynb ● + ■ + X ① Ď ▶ ■ C ▶ Code ∨ ③ sit	□ + ☆ □ □ C → Code ~ ③ git ¥ Python 3 (ipykernel)
[1]: # default_exp card	- Deck
Card Define a card in a deck of cards	Playing Cards
<pre>[2]: #lexport from fastcore.all import * import random</pre>	<pre>[] # export from deck_of_cards.card import * from fastcore.basics import * from fastcore.test import *</pre>
<pre>[3]: #[export suits = ["A", "\$", "\$", "\$"] ranks = [None, "A"] + [str(x) for x in range(2,11)] + ["J", "0", "K"]</pre>	import random
<pre>[4]: #[export class Card: "Represents a standard playing card." definit(self, suit, rank): self.suit_ent, rank = suit, rank self.suit_nm,self.rank_nm = suits[self.suit],ranks[self.rank] defeq_(self, other): return (self.suit, self.rank) == (other.suit, other.rank) deflt(self, other): return (self.suit, self.rank) < (other.suit, other.rank) deflt(self, other): return (self.suit, self.rank) < (other.suit, other.rank) defst(self): return fi{self.rank_nm}{self.suit_nm}' repr_ =str</pre>	<pre>[]: # export class Deck: "Represents a deck of cards" definit(self): self.cards = [Card(s, r) for s in range(4) for r in range(defstr(self): return '; '.join(map(str, self.cards)) deflen(self): return len(self.cards) defcontains(self, card): return card in self.cards repr =str</pre>
Cand, is a class that represents a single card in a deck of cards. For example:	<pre>def add(self,</pre>
[5]: Card(suit=2, rank=11) [5]: J•	"Adds `card` to the deck"
Creating Cards	
We can create carde like this:	

3. Documentation w/Quarto



<_of_cards		Q ≡ Q
eck of Cards ard eck	Card Define a card in a deck of cards	On this page Creating Cards Comparing Cards
	Card Card (suit=0, rank=2) Represents a standard playing card. Card is a class that represents a single card in a deck of cards. For example:	O Report an issue
	Card(suit=2, rank=11) J• Creating Cards We can create cards like this:	
	<pre>c = Card(suit=1, rank=3) c 3 of Diamonds</pre>	

4. README with Notebooks, Too



∃ README.md

Deck of Cards

This repo uses code from Allen Downey's ThinkPython2.

Install

After cloning this repository:

pip install -e .

There is already a project called **deck_of_cards on pypi**. This project has no relation to that. This project is an example of how to create python packages with **nbdev**.

How to use

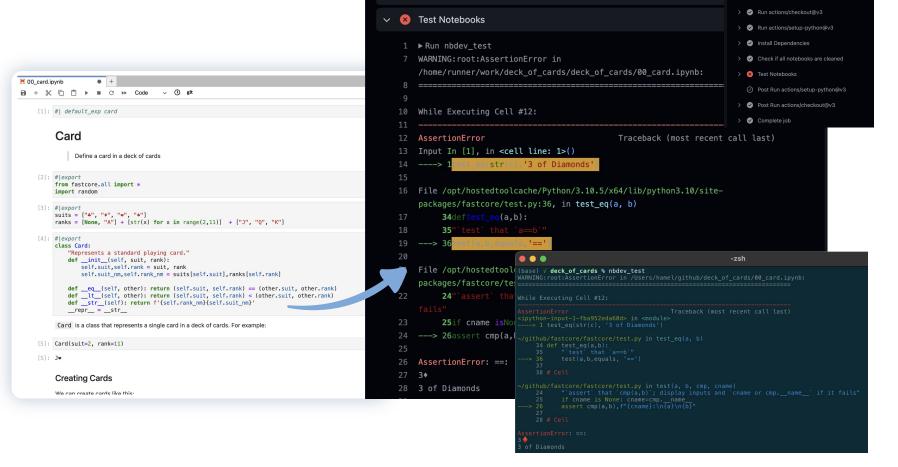
Playing cards in python!

from deck_of_cards.deck import Deck
d = Deck()
print(f'Number of playing cards in the deck: {len(d.cards)}')

Number of playing cards in the deck: 52



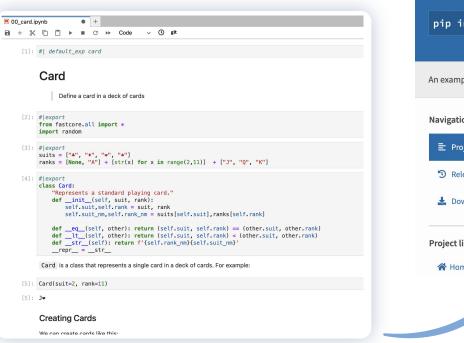
5. Tests: CI & Local



test

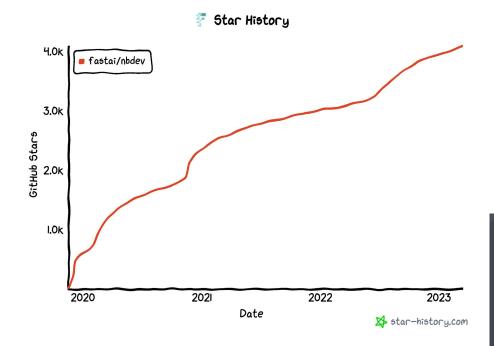
> 🧭 Set up job

6. Python Packaging



Search proje	cts Q	Help Sr		
hamel-deck-of-cards 0.0.1 pip install hamel-deck-of-cards				
An example repository for nbdev	,			
Navigation	Project description			
■ Project description	Deck of Cards			
Release history	This repo uses code from <u>Allen Downey's ThinkPython2</u> .			
🛓 Download files	Install			
Project links	After cloning this repository:			
😭 Homepage	pip install -e .			

Results



🛱 fastai Public
The fastai deep learning library
● Jupyter Notebook 🛣 22.6k 😵 7.3k
fastpages Public template
An easy to use blogging platform, with enhanced support for Jupyter Notebooks.
● Jupyter Notebook 🛱 3.3k 😵 760
ghapi Public A delightful and complete interface to GitHub's amazing API Jupyter Notebook ☆ 364 책 Apache-2.0 양 40 ⓒ 32 \$ 0 Updated 11 minutes ago



Chris Lattner Inventor of Swift and LLVM

I really do think [nbdev] is a huge step forward for programming environments.

Notebooks As Syntatic Sugar For Source Code

▲ erikgaas on Nov 20, 2020 | parent | context | favorite | on: Nbdev: A literate programming environment that dem...

So I use this in production at my company. It's an awesome tool. Personally when I'm coding in python I like to prototype in jupyter, copy code over, and then reimport anyway. Nbdev streamlines everything so I can write docs, tests, and code all in one place. And since the docs are just a jekyll site I can copy it to our documentation aws bucket in continuous integration. And with one command I can run all the notebook tests in CI as well.

The packaging is also really well thought out. I don't have to stress out about connecting setup.py with whatever publishing system we have. The settings.ini makes things sane and I can bump the version whenever I want.

A get a lot of skeptical looks when I say the source code is in notebooks, but that's just syntactic sugar for the raw source code. You still get to edit the raw code files and with one command sync everything with the notebooks. From my point of you it is close to a pareto improvement over traditional python library development.

▲ mloncode on Nov 20, 2020 | next [-]

Really interesting! Do you mind sharing what your company is? (I am the author of the blog post)

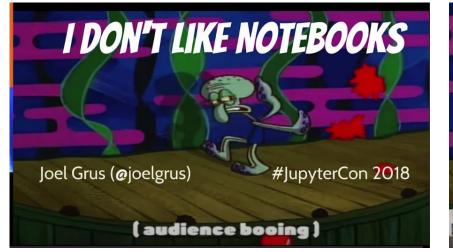
▲ erikgaas on Nov 23, 2020 | parent | next [-]

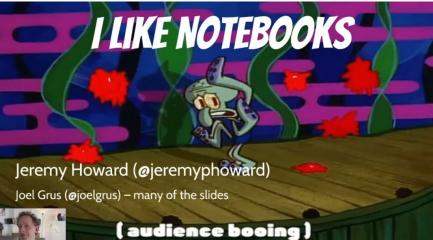
I work for Lyft's self driving car division, Level 5! Nbdev has been great. I use it a lot. Thank you for all of the work you've put into it!

Challenges With nbdev

- Source of truth isn't plain text -> diffing on GitHub
- It's effectively a new programming language in a sneaky way!
- It's bootstrapped with notebooks, rather than building an IDE from the ground up: rough edges / limitations.

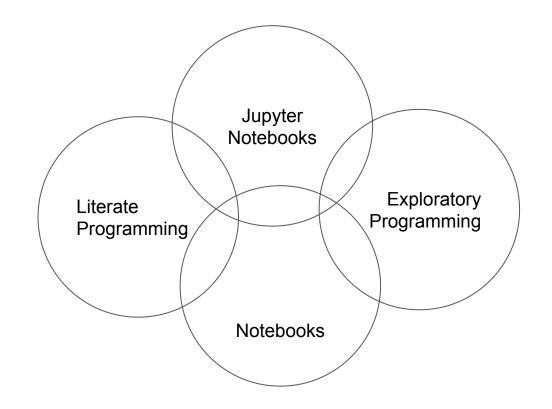
How We Write Code is Hotly Debated





We Should See The Forest For The Trees

- Literate programming and expoloratory programming are not synonymous with Jupyter notebooks.
- We can innovate on notebooks
- We can use something else
- We are empowered to make tools in this space



These Ideas Are Now More Important Than Ever

Model: Code Interpreter ALPHA



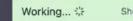
Thank you for uploading the file "plugins-demo.mp4." How can I assist you with this file?

ሪ ም



</>>

Can you extract the first 5 s of the video?



Show work 🗸

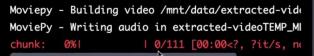
 $end_time = 5$

extracted_clip = video_clip.subclip(start_time, en

Save the extracted clip output_path = "/mnt/data/extracted-video.mp4" extracted_clip.write_videofile(output_path, codec=

output_path

STDOUT/STDERR



□ Stop generating

Stop generating

These Ideas Are Now More Important Than Ever

%%ai_ask

Can you write a function that takes a list of numbers and returns the sum of the numbers?

√ 7.3s

Sure! Here's a simple Python function that takes a list of numbers as input and returns their sum:

```
def sum_of_numbers(numbers):
    return sum(numbers)
```

This function uses the built-in sum() function to add up all the numbers in the input list. You can call this function and pass in a list of numbers like this:

my_list = [1, 2, 3, 4, 5] total = sum_of_numbers(my_list) print(total) # Output: 15

Of course, you can modify or customize this function depending on your requirements.

%%ai_continue

Can you make the retun value be a product of the numbers and not a sum?

√ 11.0s

Sure, I can easily modify the previous function to return the product of the numbers instead of their sum. Here's the updated function:

```
def product_of_numbers(numbers):
    product = 1
    for num in numbers:
        product *= num
    return product
```

This function initializes a variable called product to 1 and then uses a for loop to iterate through the input list of numbers. For each number in the list, it multiplies it with the current value of product. Finally, it returns the resulting product.

You can call this function and pass in a list of numbers like this:

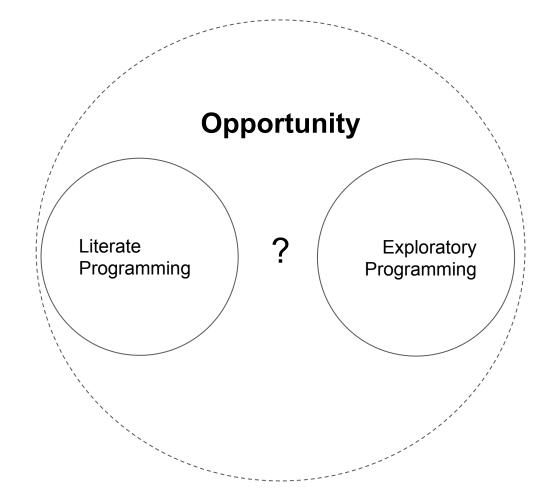
```
my_list = [1, 2, 3, 4, 5]
total = product_of_numbers(my_list)
print(total) # Output: 120
```

Again, you can modify this function according to your specific needs.

Source: Radek Osmulski @radekosmulski

Call To Action

- Let's build tools.
- Think outside the box, don't get stuck in "Notebooks for analytics" or "Hosted collaborative notebooks".
- Few people are working on this problem.



Get In Touch

https://hamel.dev