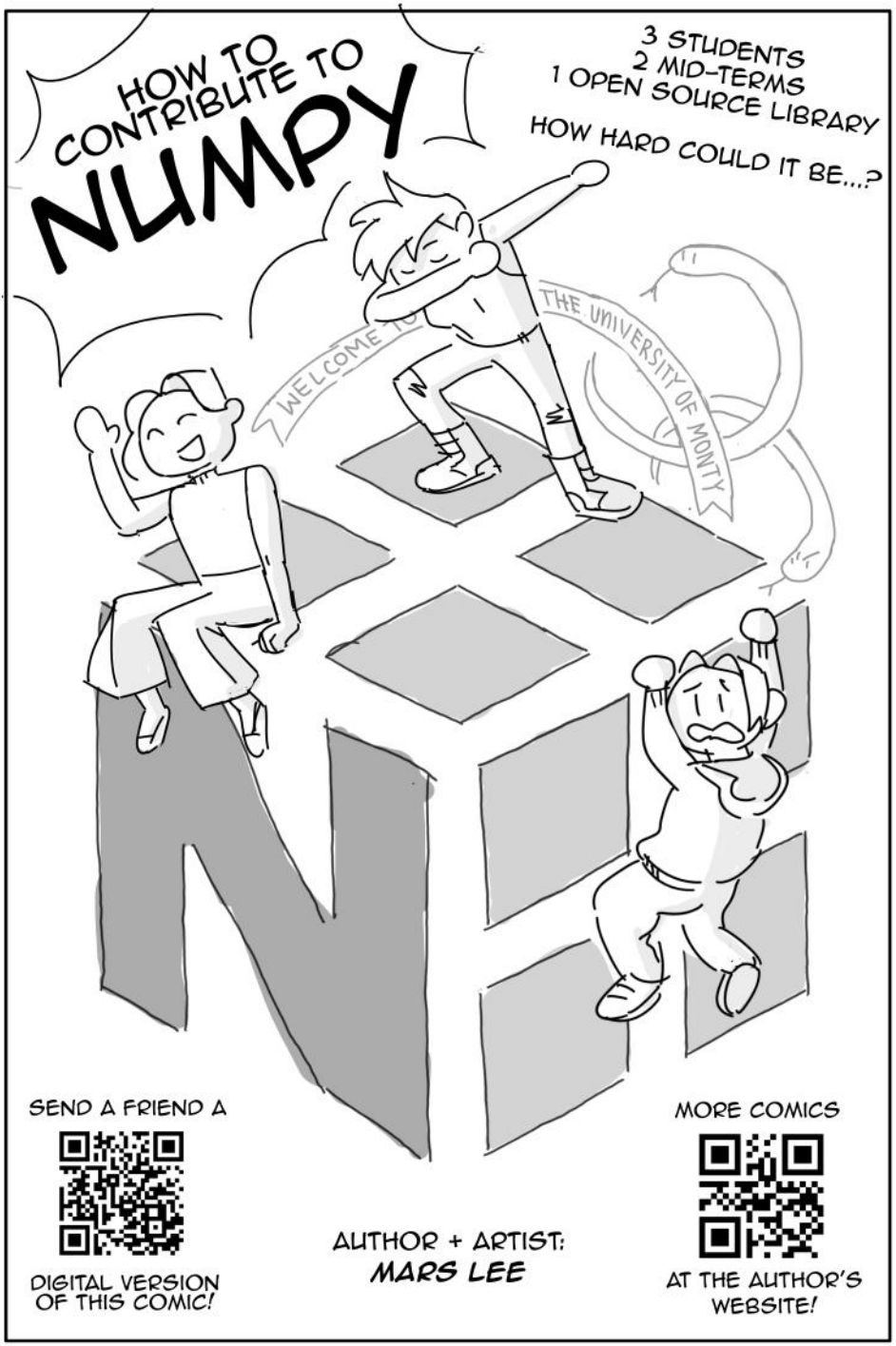
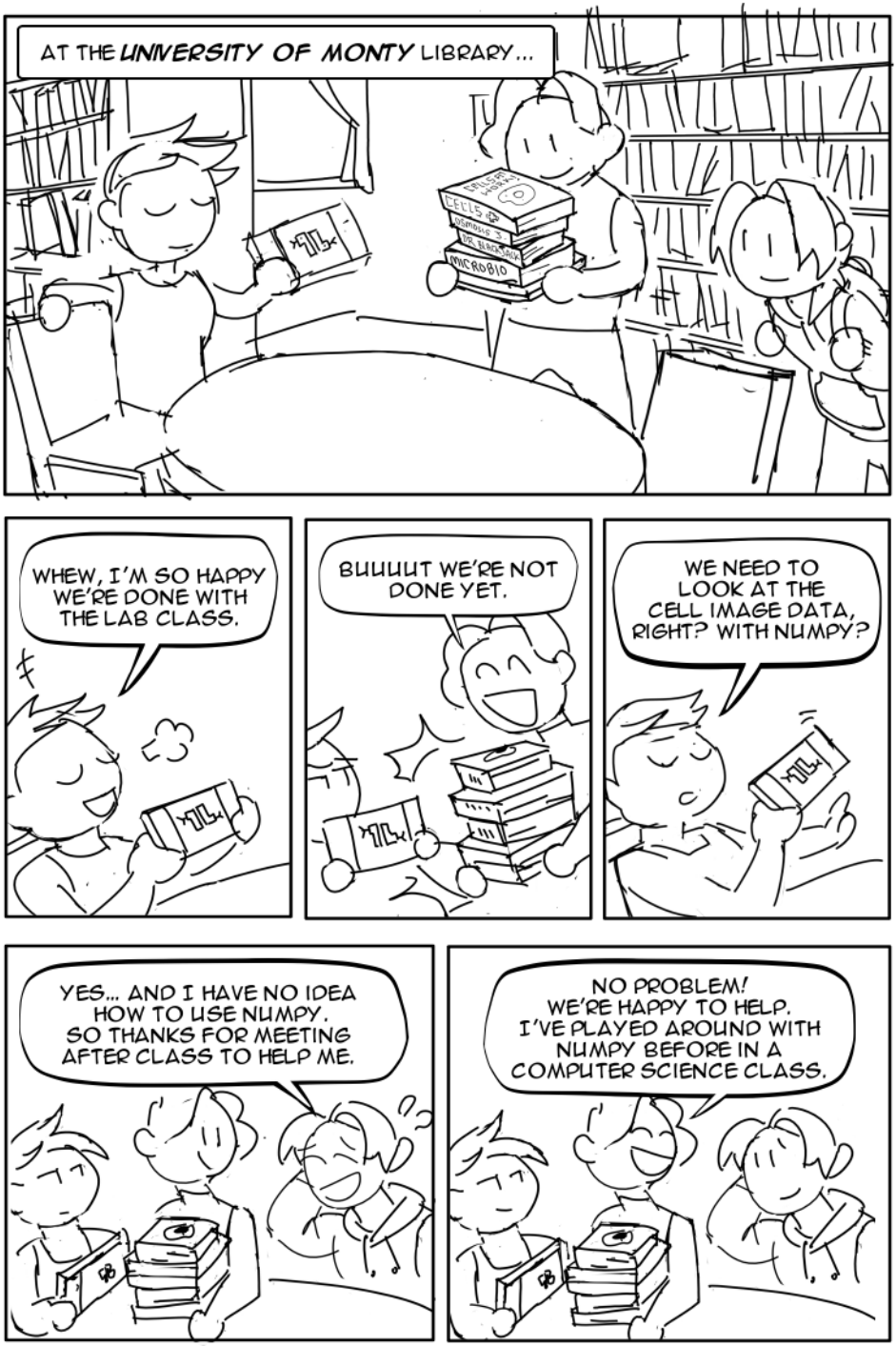
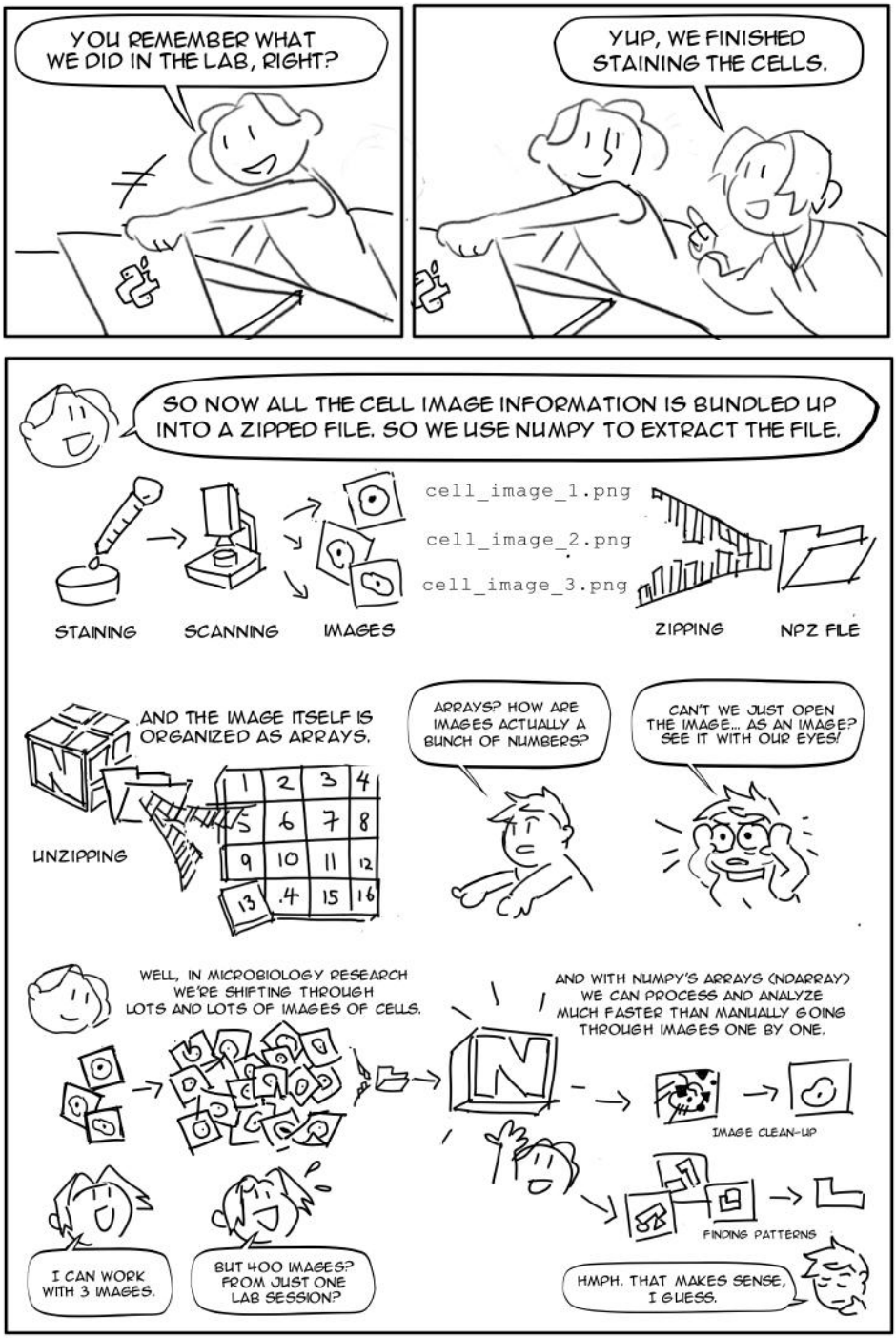
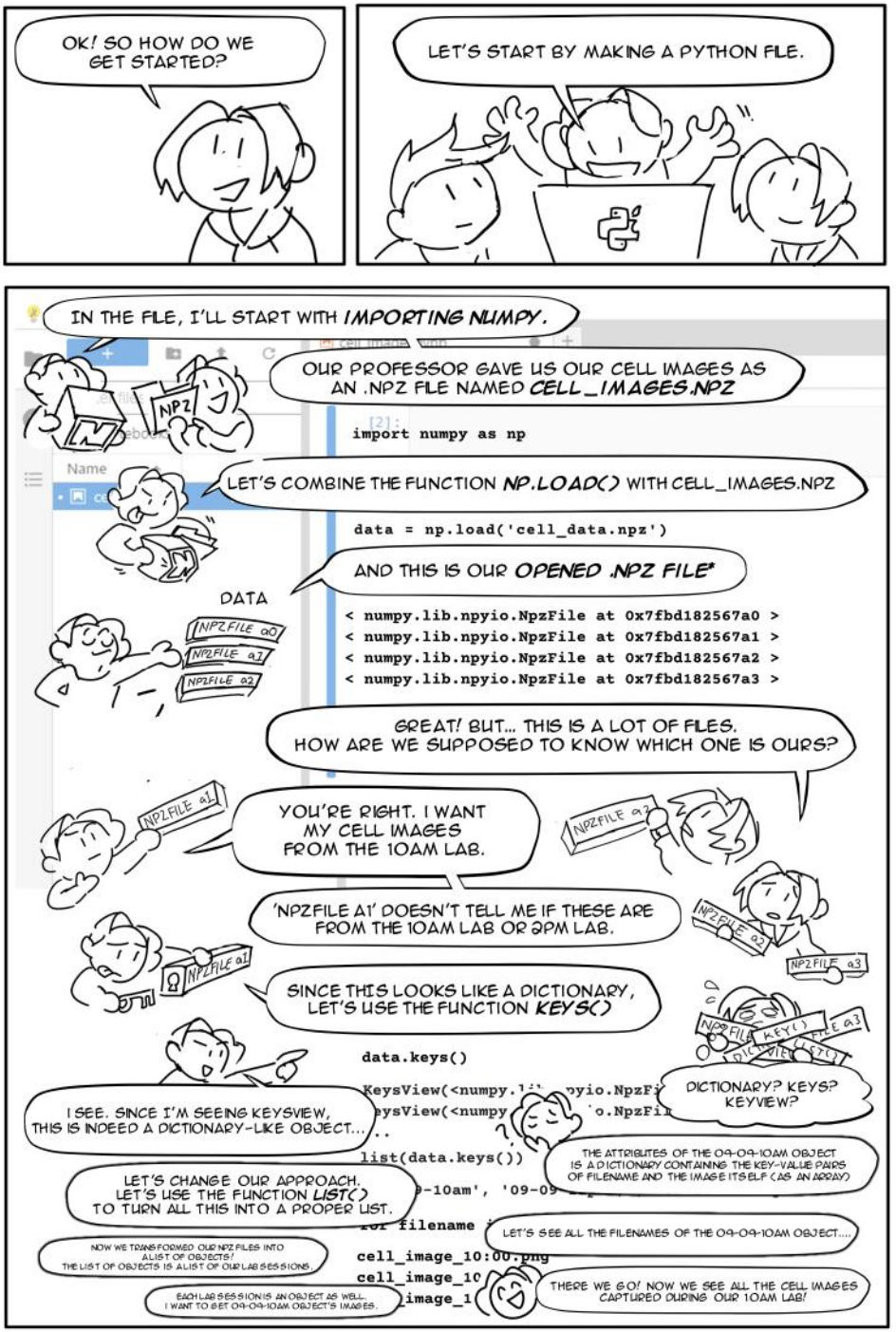

HOW TO CONTRIBUTE TO NUMPY
3 STUDENTS
2 MID-TERMS
1 OPEN SOURCE LIBRARY
HOW HARD COULD IT BE...?
WELCOME TO
THE UNIVERSITY OF MONTY
SEND A FRIEND A
DIGITAL VERSION OF THIS COMIC!
AUTHOR + ARTIST:
MARS LEE
MORE COMICS
AT THE AUTHOR'S WEBSITE!


AT THE UNIVERSITY OF MONTY LIBRARY...
CELLS
MICROBIO
WHEW, I'M SO HAPPY WE'RE DONE WITH THE LAB CLASS.
BUUUUT WE'RE NOT DONE YET.
WE NEED TO LOOK AT THE CELL IMAGE DATA, RIGHT? WITH NUMPY?
YES... AND I HAVE NO IDEA HOW TO USE NUMPY. SO THANKS FOR MEETING AFTER CLASS TO HELP ME.
NO PROBLEM! WE'RE HAPPY TO HELP. I'VE PLAYED AROUND WITH NUMPY BEFORE IN A COMPUTER SCIENCE CLASS.

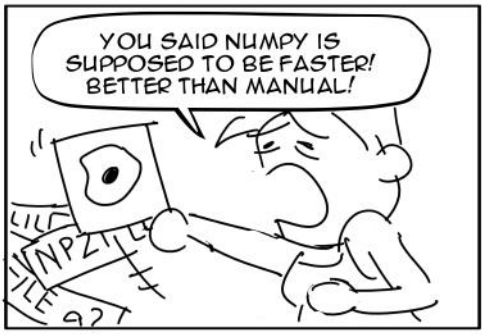

YOU REMEMBER WHAT WE DID IN THE LAB, RIGHT?
YUP, WE FINISHED STAINING THE CELLS.
SO NOW ALL THE CELL IMAGE INFORMATION IS BUNDLED UP INTO A ZIPPED FILE. SO WE USE NUMPY TO EXTRACT THE FILE.
cell_image_1.png
cell_image_2.png
cell_image_3.png
STAINING
SCANNING
IMAGES
ZIPPING
NPZ FILE
AND THE IMAGE ITSELF IS ORGANIZED AS ARRAYS.
UNZIPPING
ARRAYS? HOW ARE IMAGES ACTUALLY A BUNCH OF NUMBERS?
CAN'T WE JUST OPEN THE IMAGE... AS AN IMAGE? SEE IT WITH OUR EYES!
WELL, IN MICROBIOLOGY RESEARCH WE'RE SHIFTING THROUGH LOTS AND LOTS OF IMAGES OF CELLS.
AND WITH NUMPY'S ARRAYS (NDARRAY) WE CAN PROCESS AND ANALYZE MUCH FASTER THAN MANUALLY GOING THROUGH IMAGES ONE BY ONE.
IMAGE CLEAN-UP
FINDING PATTERNS
I CAN WORK WITH 3 IMAGES.
BUT 400 IMAGES? FROM JUST ONE LAB SESSION?
HMPH. THAT MAKES SENSE, I GUESS.

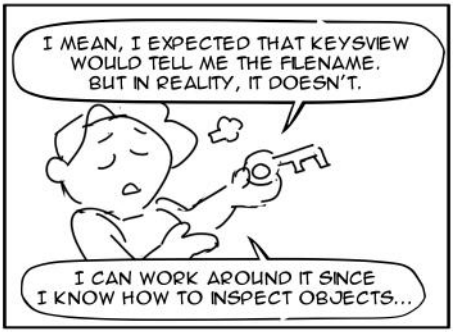


*BEFORE THE NUMPY. 1.25 RELEASE

File Edit View Run Kernel Tabs Settings Help
Name
Last Modified
cell_image.i...
2 minutes ago
WHAT-
THE-?!
NPZFILE a0
NPZFILE a2

WHY CAN'T YOU SEE THE FILENAME RIGHT THERE?
NPZFILE a0


YOU SAID NUMPY IS SUPPOSED TO BE FASTER! BETTER THAN MANUAL!

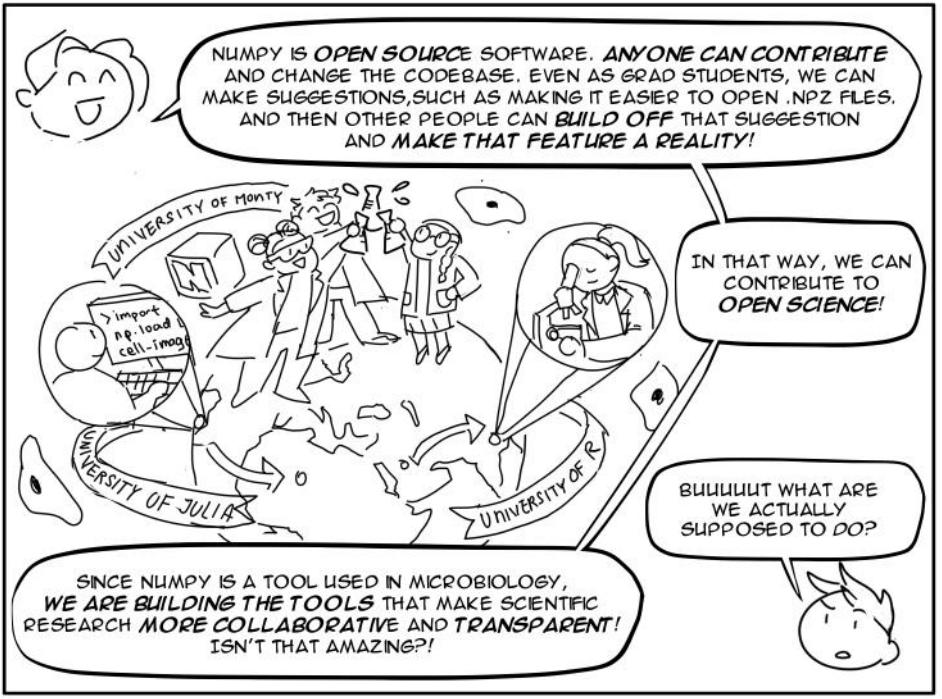

I MEAN, I EXPECTED THAT KEYSVIEW WOULD TELL ME THE FILENAME. BUT IN REALITY, IT DOESN'T.
I CAN WORK AROUND IT SINCE I KNOW HOW TO INSPECT OBJECTS...

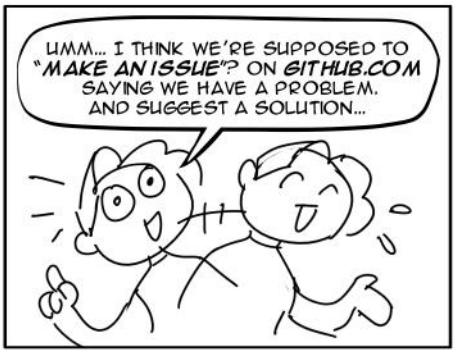
SO YOU NEED TO WORK AROUND ITS WONKINESS!

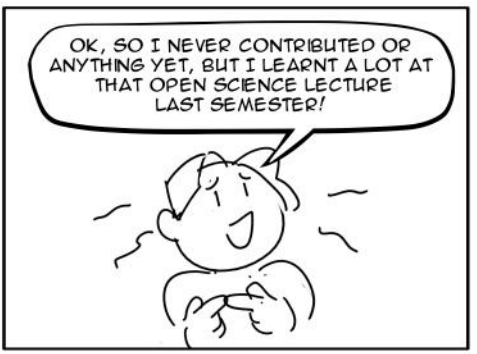
BUT, HEY, IT'S ALRIGHT, AS LONG AS IT WORKS
NPZFILE a1

I DON'T THINK WE SHOULD ACCEPT 'AS LONG AS IT WORKS' IT MAKES OUR LIVES HARDER!

YOU'RE RIGHT. IT IS HARD. BUT WE CAN CHANGE THIS!

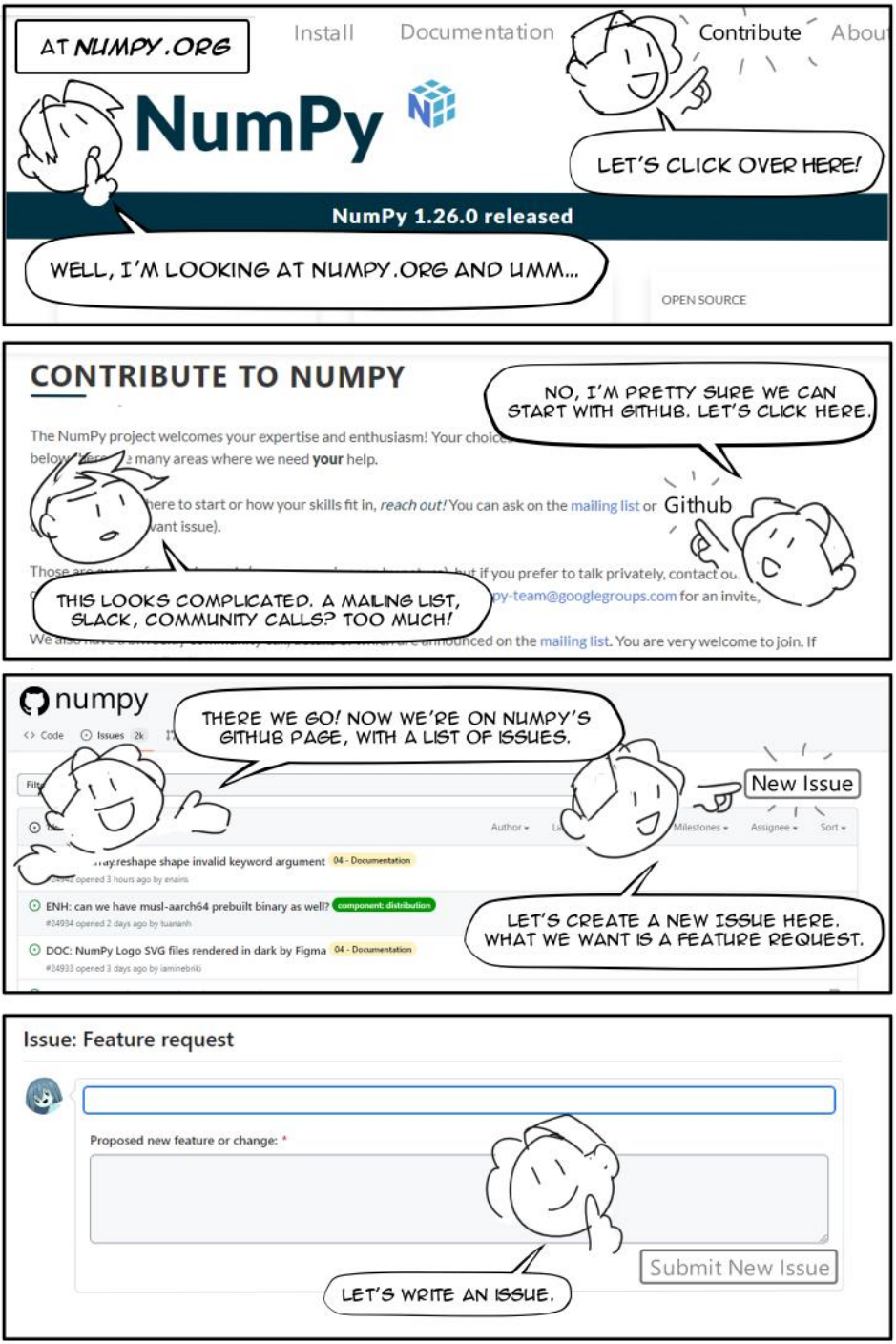

NUMPY IS OPEN SOURCE SOFTWARE. ANYONE CAN CONTRIBUTE AND CHANGE THE CODEBASE. EVEN AS GRAD STUDENTS, WE CAN MAKE SUGGESTIONS, SUCH AS MAKING IT EASIER TO OPEN .NPZ FILES. AND THEN OTHER PEOPLE CAN BUILD OFF THAT SUGGESTION AND MAKE THAT FEATURE A REALITY!
UNIVERSITY OF MONTY
>import np.load cell.imag
UNIVERSITY OF JULIA
UNIVERSITY OF R
IN THAT WAY, WE CAN CONTRIBUTE TO OPEN SCIENCE!
BUUUUUT WHAT ARE WE ACTUALLY SUPPOSED TO DO?
SINCE NUMPY IS A TOOL USED IN MICROBIOLOGY, WE ARE BUILDING THE TOOLS THAT MAKE SCIENTIFIC RESEARCH MORE COLLABORATIVE AND TRANSPARENT! ISN'T THAT AMAZING?!


UMM... I THINK WE'RE SUPPOSED TO "MAKE AN ISSUE"? ON GITHUB.COM SAYING WE HAVE A PROBLEM. AND SUGGEST A SOLUTION...


OK, SO I NEVER CONTRIBUTED OR ANYTHING YET, BUT I LEARNT A LOT AT THAT OPEN SCIENCE LECTURE LAST SEMESTER!

I HAVE A GITHUB ACCOUNT. WE CAN TRY IT WITH MINE!

GREAT! LET'S LOOK AT NUMPY.ORG TO FIND THEIR GITHUB PAGE.

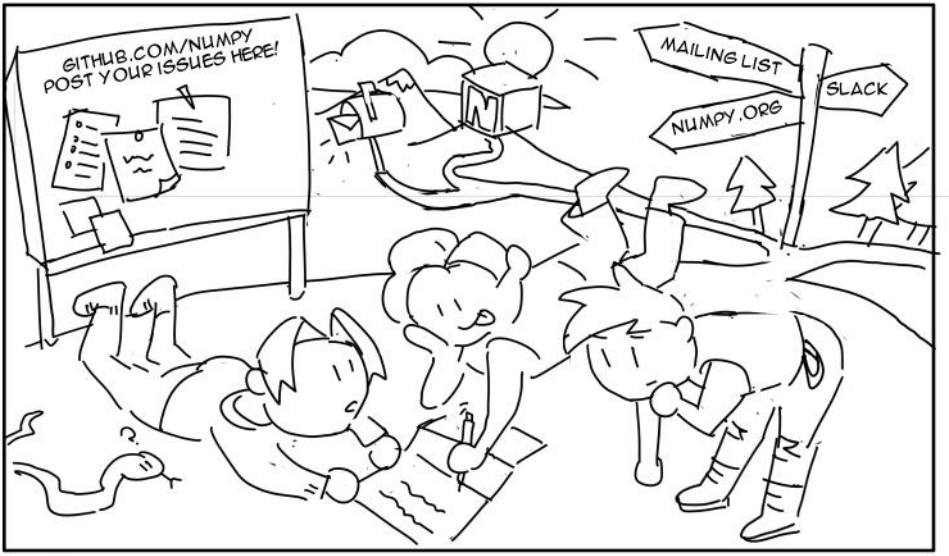

AT NUMPY.ORG
Install
Documentation
Contribute
NumPy
LET'S CLICK OVER HERE!
NumPy 1.26.0 released
WELL, I'M LOOKING AT NUMPY.ORG AND UMM...
OPEN SOURCE
CONTRIBUTE TO NUMPY
NO, I'M PRETTY SURE WE CAN START WITH GITHUB. LET'S CLICK HERE.
The NumPy project welcomes your expertise and enthusiasm! Your choice
many areas where we need your help.
here to start or how your skills fit in, reach out! You can ask on the mailing list or
Github
vant issue).
THIS LOOKS COMPLICATED. A MAILING LIST, SLACK, COMMUNITY CALLS? TOO MUCH!
if you prefer to talk privately, contact ou
py-team@googlegroups.com for an invite,
on the mailing list. You are very welcome to join. If
numpy
THERE WE GO! NOW WE'RE ON NUMPY'S GITHUB PAGE, WITH A LIST OF ISSUES.
Code
Issues 2k
New Issue
Author
Milestones
Assignee
Sort
reshape shape invalid keyword argument
04 - Documentation
ENH: can we have musl-aarch64 prebuilt binary as well?
component: distribution
#24934 opened 2 days ago by tuananh
DOC: NumPy Logo SVG files rendered in dark by Figma
04 - Documentation
#24933 opened 3 days ago by iaminebriki
LET'S CREATE A NEW ISSUE HERE. WHAT WE WANT IS A FEATURE REQUEST.
Issue: Feature request
Proposed new feature or change: *
Submit New Issue
LET'S WRITE AN ISSUE.

Issue: Feature request

Show filenames in npz archive?

Proposed new feature or change: *

Hi, we're microbiology students using NumPy to look at our cell images



PROBLEM: We found it difficult to quickly see the filename in an npz file.

Example code:

```
>>> data = np.load("example.npz")
```

I expected that KeysView would tell me the filename. But it doesn't.

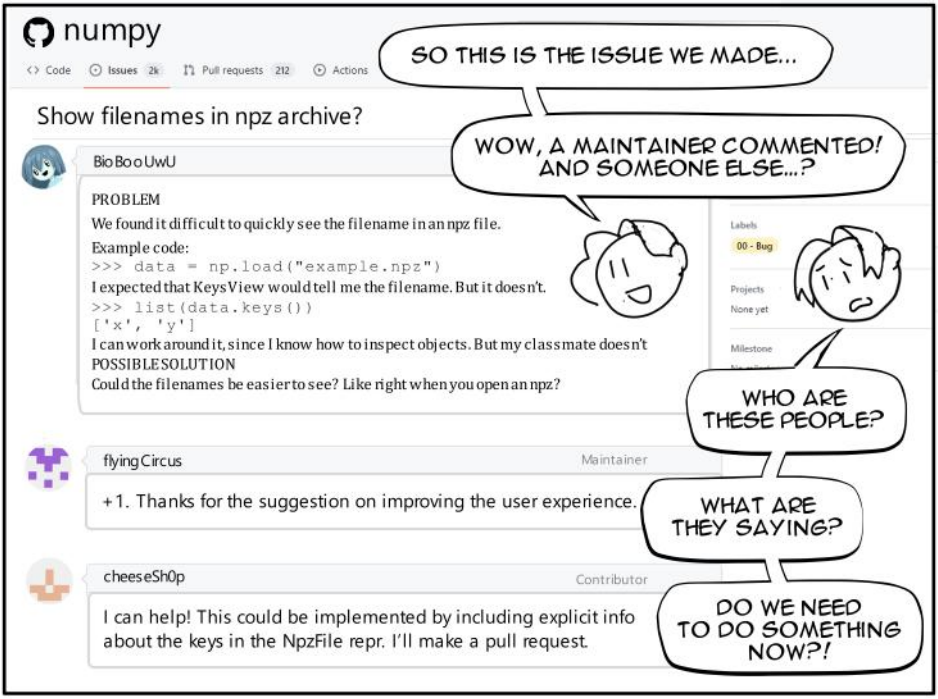
```
>>> list(data.keys())
['x', 'y']
```

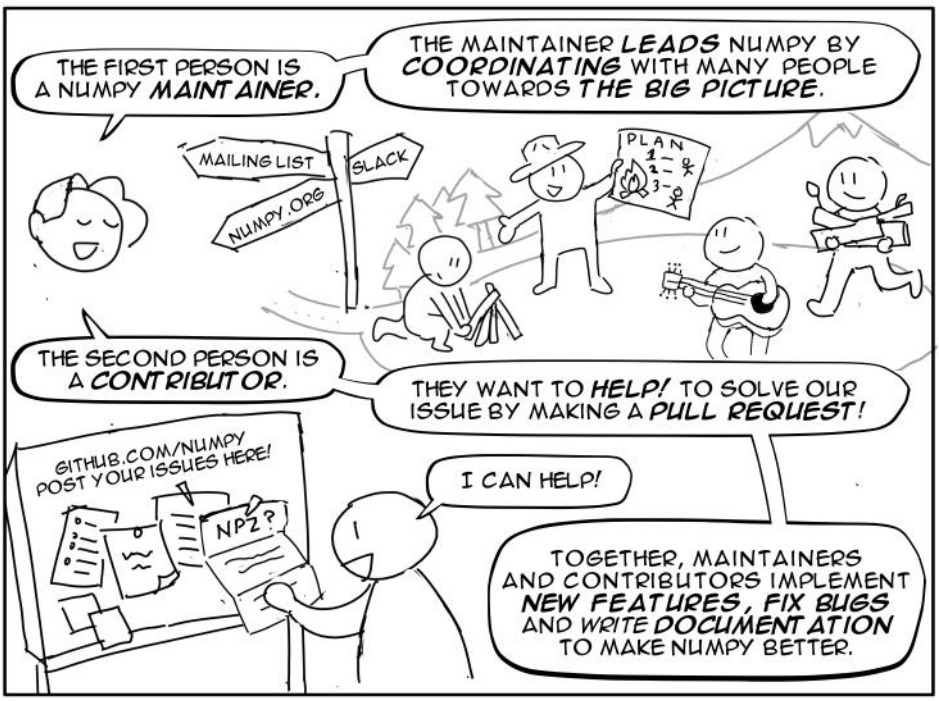
I can work around it, since I know how to inspect objects.

But my classmate doesn't.

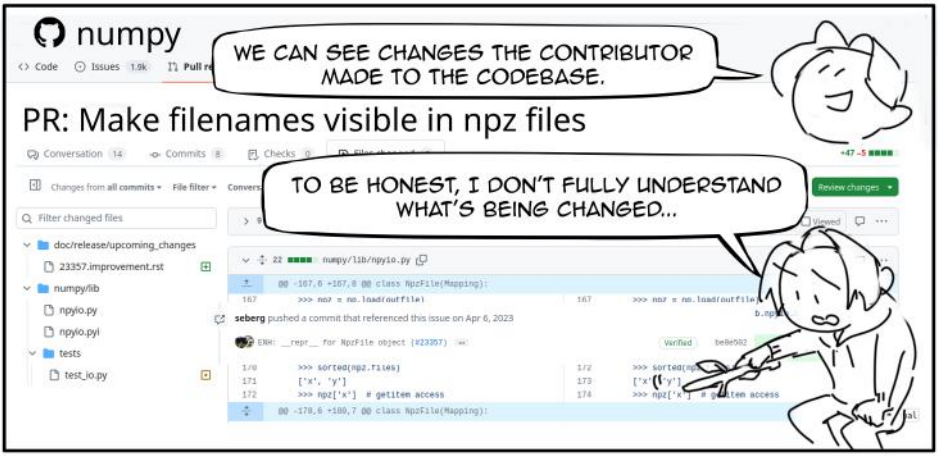
NOW TELL THEM TO SOLVE IT.
I DON'T THINK WE SHOULD BE SO DEMANDING!
HOW ABOUT WE SUGGEST A POSSIBLE SOLUTION?
POSSIBLE SOLUTION
Could the filenames be easier to see? Like right when you open an npz?
SHOULD I BE MORE SPECIFIC? I DONT KNOW HOW TO WRITE THE CODE FOR A SOLUTION...
BUT HOPEFULLY SOMEONE WILL?
yoink!
GITHUB.COM/NUMPY POST YOUR ISSUES HERE!
THUNK
NPZ?
WOO! SUBMITTED! INTO THE INTERNET!
NPZ

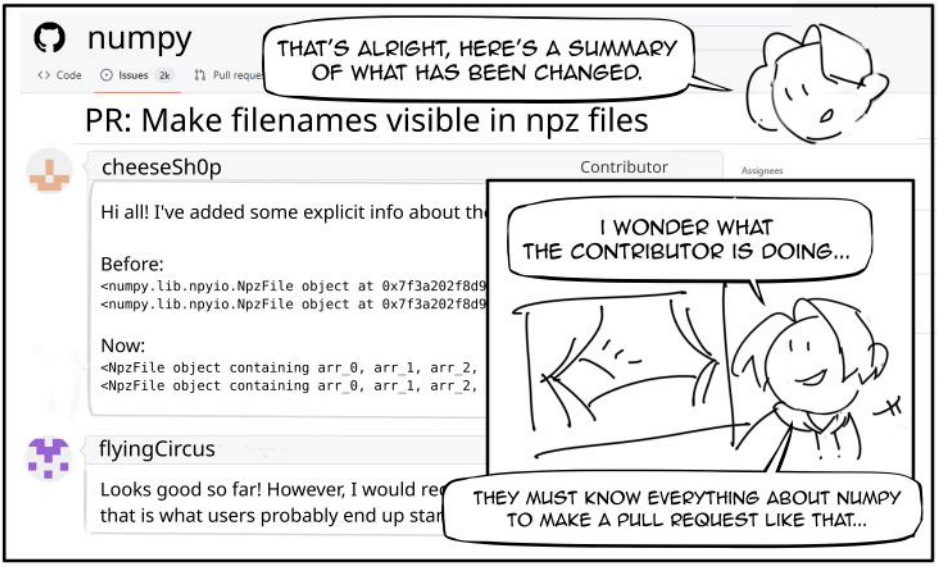
AFTER ALL THAT WORK, I NEED A BREAK.
B-BUT WE DIDN'T FINISH OUR ASSIGNMENT!
WE CAN STILL USE NUMPY, IT'S JUST ... INCONVENIENT.
WELL, I HAVE GITHUB NOTIFICATIONS ON.
SO DON'T WORRY, WE'LL HEAR BACK.
SOME TIME LATER...
OH, SOMEONE RESPONDED!!
BAM-
-WHAM-
AGHH. FINE! LET'S TAKE A LOOK.

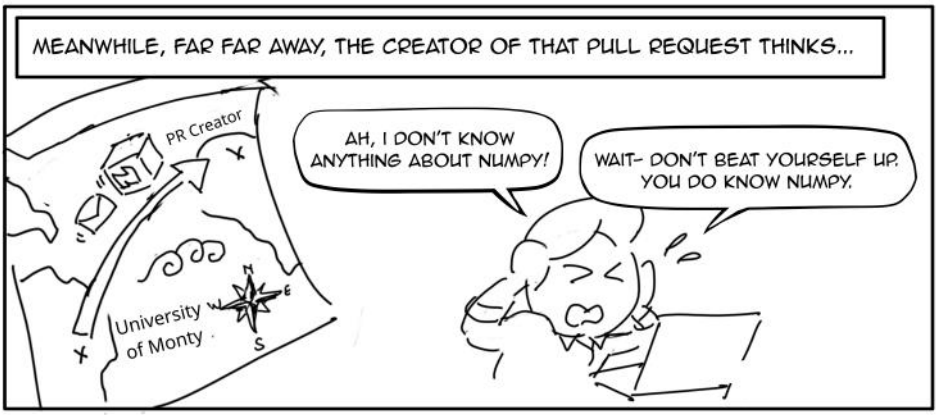

numpy
Code
Issues 2k
Pull requests 212
Actions
Show filenames in npz archive?
BioBooUwU
PROBLEM
We found it difficult to quickly see the filename in an npz file.
Example code:
>>> data = np.load("example.npz")
I expected that KeysView would tell me the filename. But it doesn't.
>>> list(data.keys())
['x', 'y']
I can work around it, since I know how to inspect objects. But my classmate doesn't
POSSIBLE SOLUTION
Could the filenames be easier to see? Like right when you open an npz?
Labels
00 - Bug
Projects
None yet
Milestone
flyingCircus
Maintainer
+1. Thanks for the suggestion on improving the user experience.
cheeseSh0p
Contributor
I can help! This could be implemented by including explicit info about the keys in the NpzFile repr. I'll make a pull request.
SO THIS IS THE ISSUE WE MADE...
WOW, A MAINTAINER COMMENTED! AND SOMEONE ELSE...?
WHO ARE THESE PEOPLE?
WHAT ARE THEY SAYING?
DO WE NEED TO DO SOMETHING NOW?!

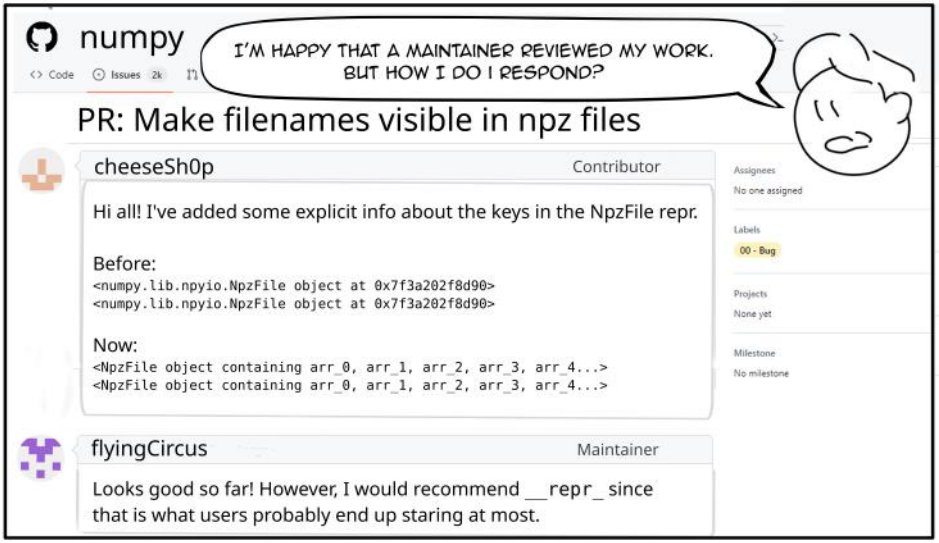

THE FIRST PERSON IS A NUMPY MAINTAINER.
THE MAINTAINER LEADS NUMPY BY COORDINATING WITH MANY PEOPLE TOWARDS THE BIG PICTURE.
MAILING LIST
SLACK
NUMPY.ORG
PLAN
THE SECOND PERSON IS A CONTRIBUTOR.
THEY WANT TO HELP! TO SOLVE OUR ISSUE BY MAKING A PULL REQUEST!
GITHUB.COM/NUMPY POST YOUR ISSUES HERE!
NPZ?
I CAN HELP!
TOGETHER, MAINTAINERS AND CONTRIBUTORS IMPLEMENT NEW FEATURES, FIX BUGS AND WRITE DOCUMENTATION TO MAKE NUMPY BETTER.

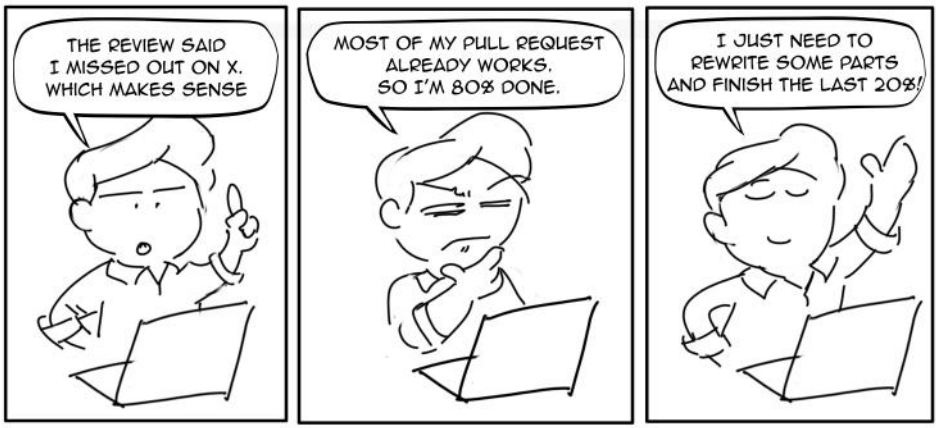
THE CONTRIBUTOR MADE A PULL REQUEST! WHAT'S THAT?
A PULL REQUEST IS ASKING TO MAKE CHANGES TO THE NUMPY CODEBASE. THIS ONE IS CHANGING THE NPZ FILES.
cheeseSh0p
Contributor
I can help! This could be implemented by including explicit info about the keys in the NpzFile repr. I'll make a pull request.
cheeseSh0p made a pull request:
PR: Make filenames visible in npz files
click

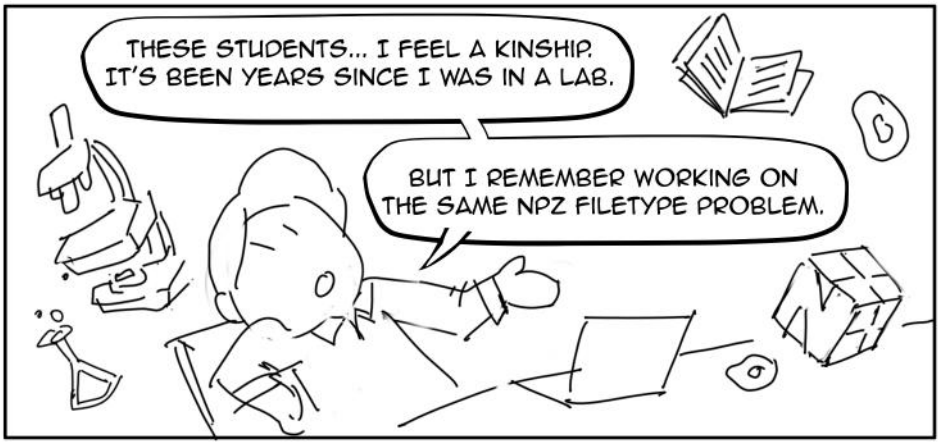

numpy
WE CAN SEE CHANGES THE CONTRIBUTOR MADE TO THE CODEBASE.
PR: Make filenames visible in npz files
TO BE HONEST, I DON'T FULLY UNDERSTAND WHAT'S BEING CHANGED...
seberg pushed a commit that referenced this issue on Apr 6, 2023

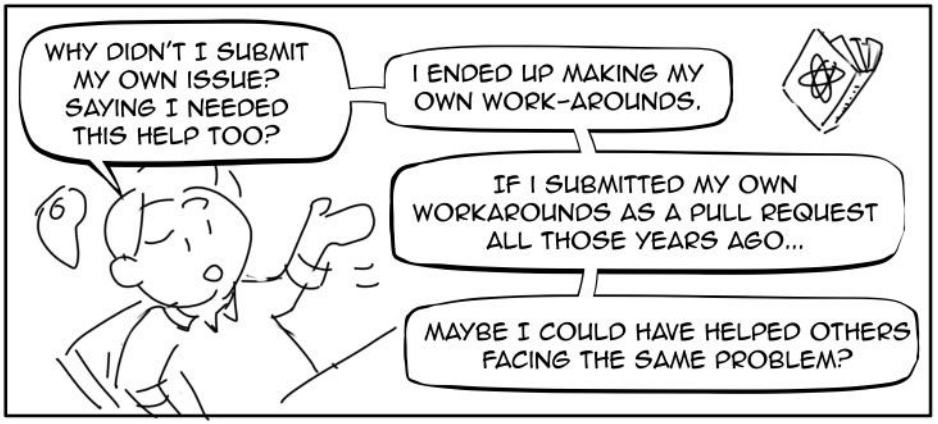

numpy
THAT'S ALRIGHT, HERE'S A SUMMARY OF WHAT HAS BEEN CHANGED.
PR: Make filenames visible in npz files
cheeseSh0p
Contributor
Assignees
Hi all! I've added some explicit info about th
Before:
<numpy.lib.npyio.NpzFile object at 0x7f3a202f8d9
<numpy.lib.npyio.NpzFile object at 0x7f3a202f8d9
Now:
<NpzFile object containing arr_0, arr_1, arr_2,
<NpzFile object containing arr_0, arr_1, arr_2,
flyingCircus
Looks good so far! However, I would re
that is what users probably end up star
I WONDER WHAT THE CONTRIBUTOR IS DOING...
THEY MUST KNOW EVERYTHING ABOUT NUMPY TO MAKE A PULL REQUEST LIKE THAT...


MEANWHILE, FAR FAR AWAY, THE CREATOR OF THAT PULL REQUEST THINKS...
PR Creator
University of Monty
AH, I DON'T KNOW ANYTHING ABOUT NUMPY!
WAIT- DON'T BEAT YOURSELF UP. YOU DO KNOW NUMPY.

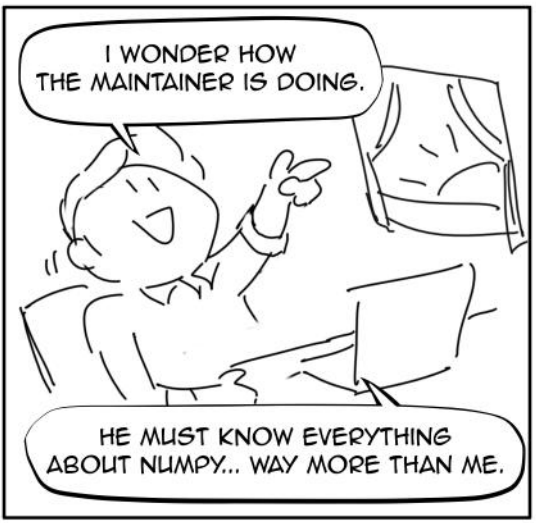

numpy
Code
Issues 2k
I'M HAPPY THAT A MAINTAINER REVIEWED MY WORK. BUT HOW I DO I RESPOND?
PR: Make filenames visible in npz files
cheeseSh0p
Contributor
Hi all! I've added some explicit info about the keys in the NpzFile repr.
Before:
<numpy.lib.npyio.NpzFile object at 0x7f3a202f8d90>
<numpy.lib.npyio.NpzFile object at 0x7f3a202f8d90>
Now:
<NpzFile object containing arr_0, arr_1, arr_2, arr_3, arr_4...>
<NpzFile object containing arr_0, arr_1, arr_2, arr_3, arr_4...>
Assignees
No one assigned
Labels
00 - Bug
Projects
None yet
Milestone
No milestone
flyingCircus
Maintainer
Looks good so far! However, I would recommend __repr_ since that is what users probably end up staring at most.

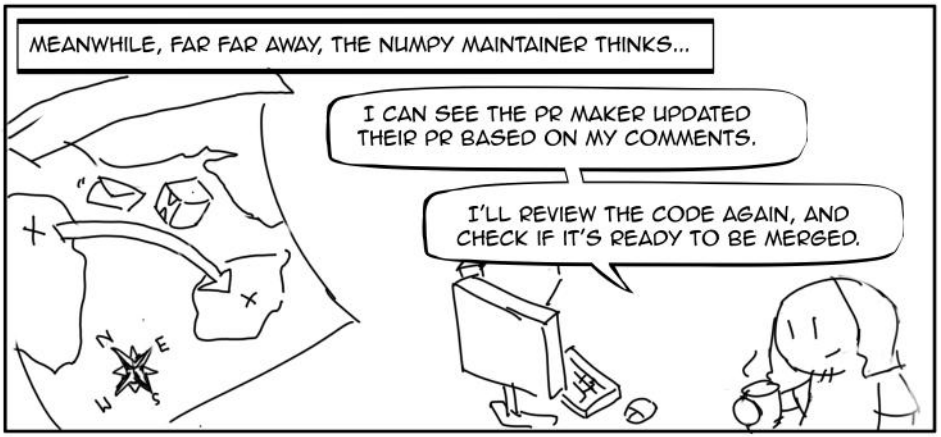

THE REVIEW SAID I MISSED OUT ON X. WHICH MAKES SENSE
MOST OF MY PULL REQUEST ALREADY WORKS. SO I'M 80% DONE.
I JUST NEED TO REWRITE SOME PARTS AND FINISH THE LAST 20%!

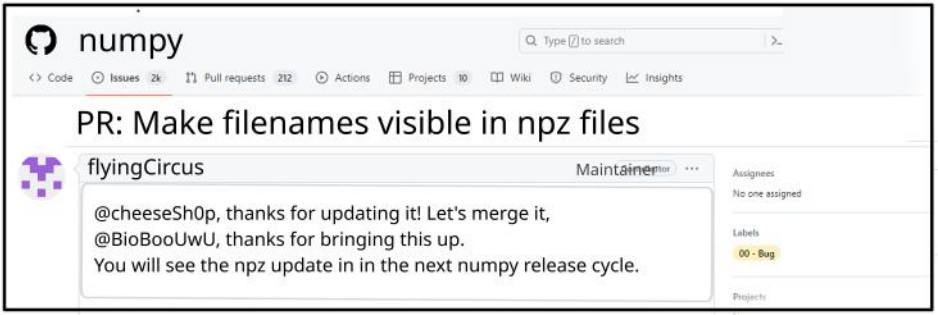

THESE STUDENTS... I FEEL A KINSHIP. IT'S BEEN YEARS SINCE I WAS IN A LAB.
BUT I REMEMBER WORKING ON THE SAME NPZ FILETYPE PROBLEM.

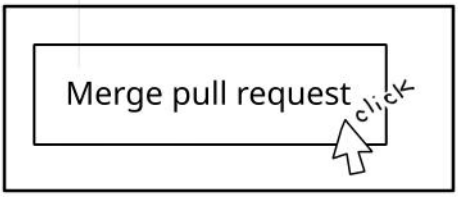

WHY DIDN'T I SUBMIT MY OWN ISSUE? SAYING I NEEDED THIS HELP TOO?
I ENDED UP MAKING MY OWN WORK-AROUNDS.
IF I SUBMITTED MY OWN WORKAROUNDS AS A PULL REQUEST ALL THOSE YEARS AGO...
MAYBE I COULD HAVE HELPED OTHERS FACING THE SAME PROBLEM?

EH! I WAS PROBABLY TO FOCUSED ON FINISHING MY THESIS TO SUBMIT A PULL REQUEST.
I'M GLAD I'M ABLE TO GIVE BACK NOW.

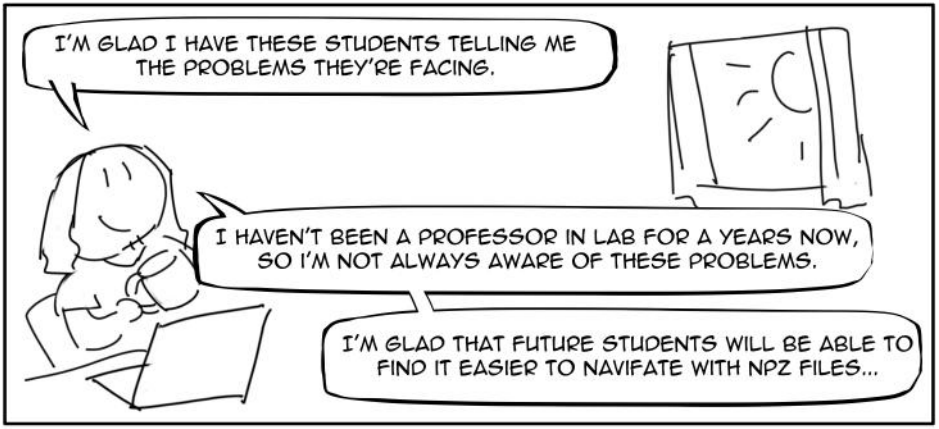

I WONDER HOW THE MAINTAINER IS DOING.
HE MUST KNOW EVERYTHING ABOUT NUMPY... WAY MORE THAN ME.

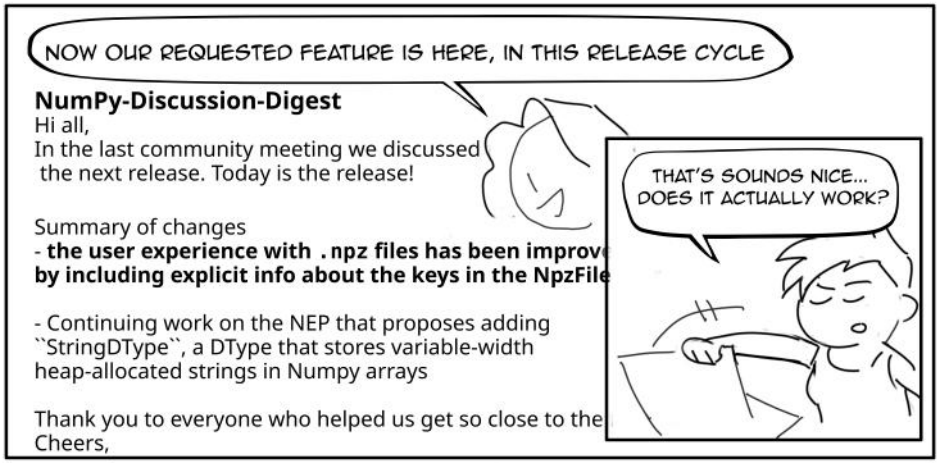

MEANWHILE, FAR FAR AWAY, THE NUMPY MAINTAINER THINKS...
I CAN SEE THE PR MAKER UPDATED THEIR PR BASED ON MY COMMENTS.
I'LL REVIEW THE CODE AGAIN, AND CHECK IF IT'S READY TO BE MERGED.


numpy
Type / to search
Code
Issues 2k
Pull requests 212
Actions
Projects 10
Wiki
Security
Insights
PR: Make filenames visible in npz files
flyingCircus
Maintainer
@cheeseSh0p, thanks for updating it! Let's merge it,
@BioBooUwU, thanks for bringing this up.
You will see the npz update in in the next numpy release cycle.
Assignees
No one assigned
Labels
00 - Bug
Projects

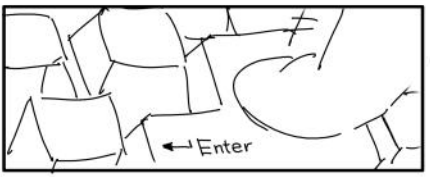

Merge pull request
click

DONE

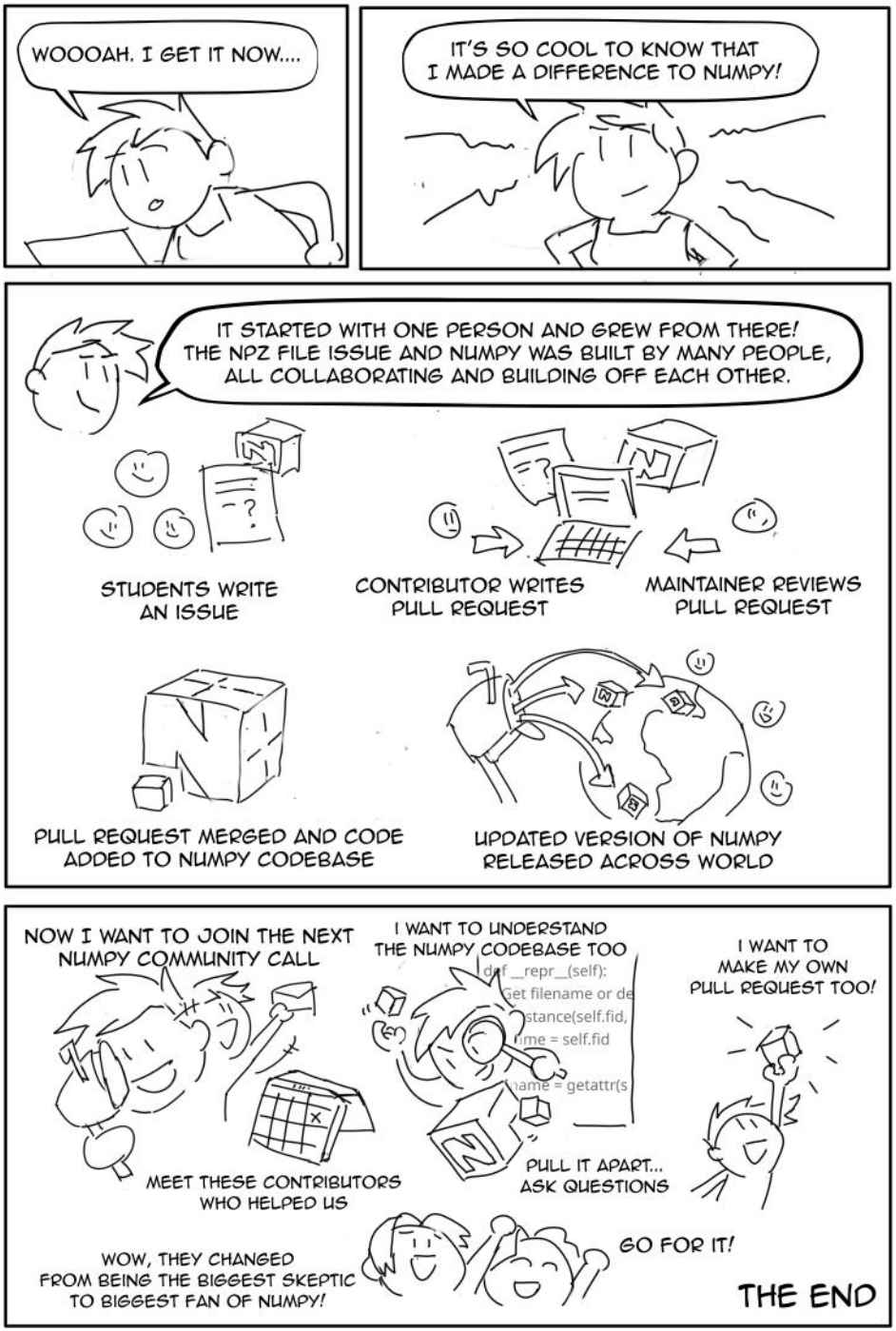

I'M GLAD I HAVE THESE STUDENTS TELLING ME THE PROBLEMS THEY'RE FACING.
I HAVEN'T BEEN A PROFESSOR IN LAB FOR A YEARS NOW, SO I'M NOT ALWAYS AWARE OF THESE PROBLEMS.
I'M GLAD THAT FUTURE STUDENTS WILL BE ABLE TO FIND IT EASIER TO NAVIFATE WITH NPZ FILES...




```
import numpy as np
data = np.load('cell)data.npz')
print(data)
```




```
<NpzFile object containing 10AM_lab_01, 10AM_lab_02, 10AM_lab
<NpzFile object containing 2PM_lab_01, 2PM_lab_02, 2PM_lab_03,
```


WOOOAH. I GET IT NOW....
IT'S SO COOL TO KNOW THAT I MADE A DIFFERENCE TO NUMPY!
IT STARTED WITH ONE PERSON AND GREW FROM THERE! THE NPZ FILE ISSUE AND NUMPY WAS BUILT BY MANY PEOPLE, ALL COLLABORATING AND BUILDING OFF EACH OTHER.
STUDENTS WRITE AN ISSUE
CONTRIBUTOR WRITES PULL REQUEST
MAINTAINER REVIEWS PULL REQUEST
PULL REQUEST MERGED AND CODE ADDED TO NUMPY CODEBASE
UPDATED VERSION OF NUMPY RELEASED ACROSS WORLD
NOW I WANT TO JOIN THE NEXT NUMPY COMMUNITY CALL
I WANT TO UNDERSTAND THE NUMPY CODEBASE TOO
def __repr__(self):
Get filename or de
stance(self.fid,
me = self.fid
name = getattr(s
I WANT TO MAKE MY OWN PULL REQUEST TOO!
MEET THESE CONTRIBUTORS WHO HELPED US
PULL IT APART... ASK QUESTIONS
GO FOR IT!
WOW, THEY CHANGED FROM BEING THE BIGGEST SKEPTIC TO BIGGEST FAN OF NUMPY!
THE END

THANKS FOR READING! I HOPE THIS COMIC INSPIRES YOU TO...

START YOUR OWN OPEN SOURCE JOURNEY!

NUMPY.ORG

NUMPY GITHUB

GOT *FEEDBACK*?
OR A *FUTURE COMIC IDEA*?
FILL OUT THE *SURVEY*!

DID YOU KNOW THIS COMIC IS

BASED ON A TRUE STORY?

IN 2023, THERE REALLY WAS AN UPDATE TO THE NPZ FILE TYPE IN NUMPY.

ROSS BARNOWSKI WAS A PROFESSOR TEACHING BIOLOGY STUDENTS. THE STUDENTS HAD DIFFICULTY ACCESSING THEIR CELL IMAGE – JUST LIKE THE COMIC!

ROSS WROTE AN ISSUE.

GANESH KATHIRESAN WROTE A PULL REQUEST CHANGING THE BEHAVIOR OF NPZ FILETYPE, FROM DISPLAYING THE MEMORY LOCATION OF THE FILE TO DISPLAYING THE FILES THEMSELVES.

MAINTAINERS ***CHUCK HARRIS*** AND ***SEBASTIAN BERG*** REVIEWED AND MERGED THE PULL REQUEST GANESH MADE.

THANK YOU:

- GOOGLE SEASON OF DOCS, QUANSIGHT AND NUMFOCUS FOR THE GRANT THAT KICKSTARTED THESE COMICS
- ***MUKULIKA PAHARI*** AND ROSS BARNOWSKI FOR BEING AWESOME PROJECT MENTORS
- THE NUMPY COMMUNITY FOR THEIR CONTINUED SUPPORT

SUBSCRIBE TO THE MARS COMIC UNIVERSE

@MARSBARLEE

LINKEDIN.COM/IN/MARS-LEE/

MARSBARLEE.COM

LIBREPUNK.CLUB/@MARSBARLEE