



# My Dinner with Numeric, NumPy, and SciPy

A Retrospective from 2001 to 2025 with Commentary and  
Anecdotes

Charles R. Harris

# In the Beginning: 2001

Music: Also sprach Zarathustra

# Who was I in 2001?

## **Not a graduate student!**

I worked at Space Dynamics Laboratory (SDL). I was hired back for my familiarity with Michelson Interferometers.

I reported to the president, Alan Steed, who gave me a great deal of freedom in choosing my tools and methods. Not every company can afford an oddball, but I was one, and a well managed one at that. Hence:

- Linux,
- Python,
- Numeric.

Only one other person used those tools

# The State of Python in 2001

- Python 1.5.2 was still commonly used.
- Python 2.0.0 came out in 2000.
- Python 2.1.0 came out in 2001.
- *Python Essential Reference* by David Beazley was published in 2000 and provided an outstanding reference for learning and extending the Python language.
- The programming environment was basic. Fernando Perez would fix that with IPython.

Documentation: excellent

# The State of Numeric in 2001

- Numeric was maintained by Paul Dubois and quite usable in 2001.
- The *Numerical Python* manual, written by David Ascher, Paul F. Dubois, Konrad Hinsen, Jim Hugunin, and Travis Oliphant, was excellent.
- The random number module was basic and buggy. My first contribution was fixing a bug in that module. Robert Kern would later provide the NumPy random module.
- Plotting was a pain to set up and use. John Hunter would fix that with Matplotlib.

Documentation: excellent.

# The State of SciPy in 2001

Created by Pearu Peterson, Eric Jones, and Travis Oliphant in 2001.

I did not find it useful at that time. But that could be fixed!

My first contributions were:

- 1D zero finders written in C, Travis cleaned it up a bit and put it in.
- The scipy constants module, posted on the mailing list.

A big benefit of contributing to an open source project is that you can fix and add things. That is especially true in the early stages of a project.

Documentation: basic.

# The State of F2Py in 2001

Pearu Peterson released the first public version of f2py in 2000. It was to become an essential element in the development of SciPy, providing the interface glue between numeric/numpy arrays and Fortran programs.

Pyrex (2002), Swig (2006) and Cython (2007) would later provide similar interfaces for C/C++ code.

Documentation: OK

Where are the Pictures?



# Python for Scientific Computing Workshop (2003)

Caltech, Guggenheim Laboratory, seating capacity 50



# Lost in the Fog and Breakfast

Fog, fog, fog, visibility is about 20 feet.

Is that a building I see before me?

Konrad Hinsen appears, “Where is it?” he asks

We eventually found the building.

Breakfast was great: Bagels, cream cheese, and a mountain of lox on ice.

There was a planning meeting after the conference, Fernando Perez was running at full speed, as he does.

# A Word From our Sponsors

## [The National Biomedical Computation Resource](#)

NBCR, UCSD, San Diego, CA

## [The Center for Advanced Computing Research](#)

CACR, CalTech, Pasadena, CA

## [Enthought, Inc.](#)

Austin, TX – Eric Jones

# Scipy-2004: The First Real Scipy

Jim Hugunin Keynote Address

## **Progress:**

- John Hunter presented Matplotlib.
- Fernando Perez presented an EMACS plugin.
- Francesc Alted presented PyTables (HDF5 interface).
- Aric Hagberg presented NetworkX.
- Perry Greenfield reported on Numarray progress.

There was steak and baked potato for dinner. The Athenaeum has a dress code?

# Stuck at the Crossroads: Numeric or Numarray

I know, I know, we should ask directions. Anyway ...

- Numarray was better for big arrays, but slower for small arrays. What to do?
- Spreading our small pool of talent over two projects threatened progress.
- Example: I wrote new sorting code and lexsort for Numarray not Numeric.

Travis Oliphant was motivated to solve the problem by merging the two.

# New Digs: Scipy-2005

Beckman Institute Auditorium, seating capacity 200



# Scipy-2005: NumPy Begins

Travis presents `scipy.base`, the NumPy progenitor.

It is still part of SciPy at this point.

# What's in a Name?

Travis Oliphant separated `scipy.base` from SciPy in 2006 and made it a separate package. The question arose, what to call it?

After much discussion we settled on Numerix, Asterix's educated cousin. But ... that name had been taken by an Australian company.

Hey, why not call it NumPy? That name was already an informal alias for Numeric, let's use that.

NumPy was born.

Don't ask me how to pronounce it :)



# What was in NumPy

NumPy was a mix of Numeric and parts of Scipy.

- F2Py (new, Pearu Peterson),
- Distutils (new, Pearu Peterson and Travis Oliphant),
- Polynomial (new, Travis Oliphant),
- Random (redone and extended, Robert Kern),
- Basic Linalg (mostly inherited),
- FFT (inherited),
- Masked arrays (inherited).

That is pretty much what NumPy still offers today.

# The Numpydoc Format Finalized December 2007

It took a year to get there, but get there we did. Pauli Virtanen (one man army) wrote the Sphinx extension for the new format and we were off.

# The NumPy Documentation Marathon (2008)

Documentation received some funding from UCF through Joe Harrington.

Pauli Virtanen set up an online documentation editor, Pydocweb.

The documentation server was initially hosted by Gaël Varoquaux's girlfriend.

A large part of NumPy's documentation was created in this period.

# Blah, Blah is Documentation?

My big contribution to the documentation was the term “array\_like”.

I also added helpful “Blah, Blah” documentation to some functions I intended to document.

Fernando Perez demonstrated NumPy’s documentation in a presentation ...

There are many ways to motivate a person :)

# Memories of Scipy-2008

It was the first Scipy to publish a proceedings. We needed to be selective about presentations and reviewing was work.

Wandering around the Caltech Campus at night looking for a room where we could work on documentation.

Robert Kern – previously at Caltech – was our restaurant guide.

We debated who would graduate first, Stefan van der Walt or David Cournapeau. The consensus was Stefan. We were wrong.

# The Gang's All Here – Scipy 2009

Many NumPy/SciPy developers from around the world attended:

- Pauli Virtanen (Finland),
- David Cournapeau (Japan),
- Stefan van der Walt (South Africa),
- Gael Varoquaux (France),
- Dag Sverre Seljebotn (Norway),
- Anne Archibald (Canada).

Python 3.0 was out, there are maybe five users in attendance. We needed to start working on support.

There were great evening conversations in the hotel corridors.

# The Three Musketeers

Travis Oliphant became President of Enthought in 2008 and faded from NumPy development after version 1.2.0. The top five committers from 1.2 to 1.5 were:

- 1179 David Cournapeau,
- 467 Pauli Virtanen,
- 386 Charles Harris,
- 116 Pierre GM,
- 86 Travis Oliphant.

The first three are the Musketeers. There was no formal transfer of power.

# Adventures of the Three Musketeers

We did the following:

- Made NumPy and F2Py Python 3 compatible (all),
- Broke up the files inherited from Numeric into functional units (all),
- Wrote new Polynomial classes (Charles Harris),
- Integrated a new masked array implementation (Pierre GM),
- Integrated Nose testing (Alan McIntyre).

Ralf Gommers replaced David Cournapeau as release manager for 1.5.



# New Digs for Scipy-2010: AT&T Conference Center



# Development Moves from Subversion to GitHub (2010)

The move was decided on in a meeting at SciPy 2010. We weren't the first project to move, IPython had moved earlier, but we wanted good Windows support before moving. Pauli Virtanen did the conversion from Subversion to Git, and Ralf Gommers released NumPy 1.5 using the GitHub source. We called him up and got an OK before making the decision.

The move significantly improved the development workflow.

## Burnout, Mark Wiebe to the Rescue

After NumPy 1.5 was out, the Three Musketeers disbanded. I felt overwhelmed doing my part to keep NumPy going, and then, Mark Wiebe appeared.

Mark showed up in Nov 2010, and it was instantly apparent that he was very, very good. His impact was immediate and I slept soundly for the first time in weeks.

# Things Mark did in NumPy 1.6 and 1.7

- Added the float16 dtype,
- Created einsum,
- Made datetime work,
- Added calendar functions,
- Added the “where” and “keepdims” keywords to ufuncs,
- Created a new array iterator,
- Improved type promotion,
- Made bit masks part of ndarray, think Arrow validity arrays. That was controversial and taken out before the 1.7 release.

# The NumPy Governance Document

Fernando Perez had told us many times that we needed a governance document so that we could solicit funding for NumPy development. We finally sat down to discuss it at SciPy-2015 and tasked Nathaniel Smith to write it up. By the end of 2015 initial versions of both the NumPy Governance Document and a Code of Conduct were in place.

# Binary Distribution Matters – Wheels

In the early years, Numpy was distributed as \*.exe files on Windows, \*.dmg files on Mac, and source files for Linux, maybe five files in all on SourceForge. That changed with the development of wheels.

- “The Wheel Binary Package Format 1.0” (PEP 427) came in 2013 and took care of Windows and Mac, but Linux was still stuck with source releases.
- “A Platform Tag for Portable Linux Built Distributions” (PEP 513, manylinux1) took care of the Linux problem in 2016. This was a big step forward, especially for downstream testing.

The latest NumPy release on PyPI had 50 wheels.

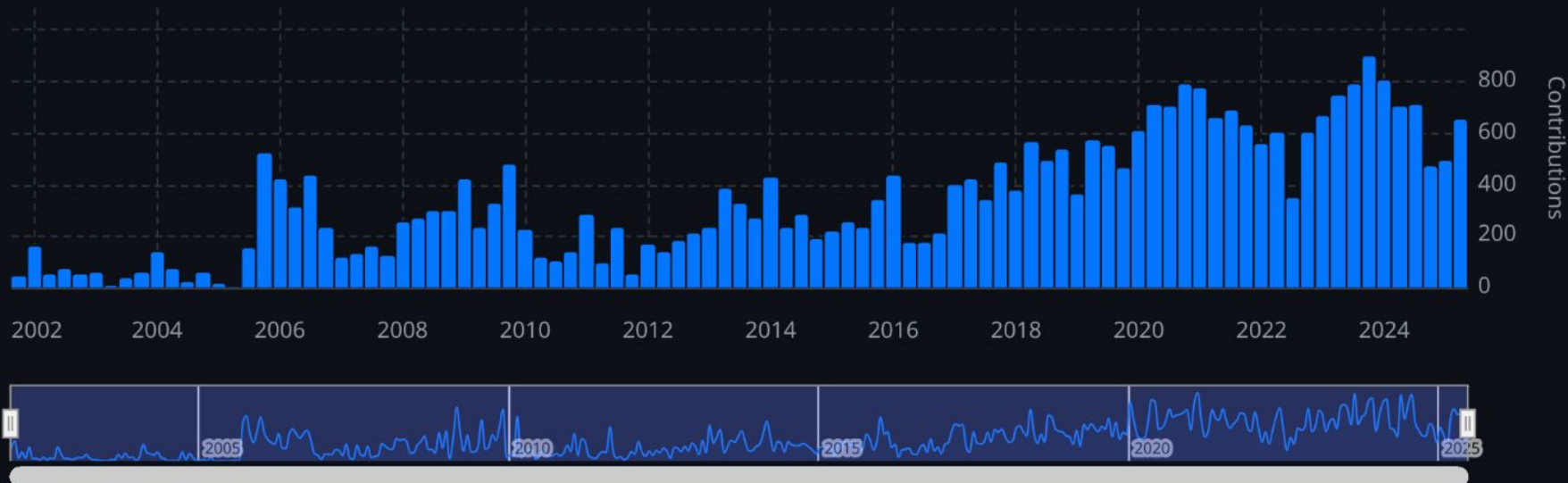
## Funded Developers at Last (2017)

In 2017 BIDS (Berkeley Institute of Data Science) received grants totaling \$1.3 million from the Moore and Sloan foundations for NumPy development. They hired two Developers, Matti Picus and Tyler Reddy. They were the first developers paid to work on NumPy, and started at the end of 2017. The rate of commits began to rise.

# Weekly Commits Before and After 2018

## Commits over time

Weekly from Dec 15, 2001 to Jun 28, 2025





# The Scientific Python Stack Diversifies

By 2018 there were many NumPy like projects. That diversification was driven by both the needs of the machine learning community and the desire to utilise the computational power offered by TPUs, GPUs, and distributed processing.

Examples are:

- Jax,
- Pytorch,
- Dask,
- CuPy,
- XArray.

As in the Numeric/Numarray split, this threatened to split the community.

# Interoperability, not Competition

NumPy did not have the resources to compete in that larger ecosystem, we decided to promote interoperability instead.

There were two initial steps in that direction:

- `__array_ufunc__` (nep-0013), allowed overriding ufuncs
- `__array_function__` (nep-0018), allowed overriding array functions

These steps were helpful, and stopped the bleeding in the short term, but the better long term solution is the Array API Standard (NEP-0056). That standard is versioned and should slowly increase its coverage in coming years.

# NumPy drops Python 2.7 (2019)

Python itself planned to drop Python 2.7 in 2020, but we needed C99 compilers on Windows and Python 2.7 was blocking that. Consequently, we dropped support a year early (NEP-0014). This allowed several important improvements:

- A new random number implementation, due to Robert Kern and Kevin Sheppard, was merged into NumPy (NEP-0019). It was a state of the art design with an eye toward parallelization.
- Replacement of FFTPACK by Pocketfft, a modern fft implementation by Martin Reinecke. This brought NumPy support up to current standards.

# NumPy moves to Meson for Building (2023)

NumPy distutils, which served NumPy well for many years, came to an end in 2023. NumPy distutils was an extension of Python distutils, and Python dropped that package in version 3.12, forcing NumPy to move away from setup based builds. That move was mostly implemented by Ralf Gommers, and took almost a year to complete, as both NumPy and Meson needed to be updated.

# Good Bye Austin, Hello Tacoma 2024



# NumPy 2.0.0 Released (2024)

In 2024 NumPy 2.0.0 was released after a year of heavy development. There were a large number of changes, the majority due to Sebastian Berg.

- Functions were updated for Array API v2022.12,
- Promotion rules were changed for Array API compatibility,
- Dtype handling was changed for improved extensibility,
- The API was changed, outdated functions and aliases were removed,
- The ABI was changed,
- Arrays of variable length strings were added (Nathan Goldbaum)

It is remarkable that NumPy is still under heavy development after all these years, but hardware and needs keep changing.

# What have we accomplished?

When I first became involved in Numeric/NumPy development, I had no idea that it would have as much impact as it has on scientific computation, all I wanted was something that was useful to me. That may be key, for NumPy development has been driven by the needs of its users. It is a service. That it is free, easily available, and light weight are also important features.

NumPy itself may fade in the future, but I believe that the Array API and the idea of universal, and free, scientific software will persist. That change in the ecology of scientific computing may end up being our greatest accomplishment.