

Interview with The Creator of Keras, AI Researcher: François Chollet

Originally published by Sanyam Bhutani

François Chollet, one of the greatest contributors of the complete community: The creator of Keras, author of one of the best Deep Learning books.

Chollet is currently working as an AI Researcher at Google, he's also at the core of Keras Development.

About the Series:

I have very recently started making some progress with my [Self-Taught Machine Learning Journey](#). But to be honest, it wouldn't be possible at all without the amazing community online and the great people that have helped me.

In this Series of Blog Posts, I talk with People that have really inspired me and whom I look up to as my role-models.

The motivation behind doing this is, you might see some patterns and hopefully you'd be able to learn from the amazing people that I have had the chance of learning from.

Sanyam Bhutani: Hello François, Thank you so much for taking the time to do this interview. It's an absolute honor to be talking with you.

François Chollet: Thanks for having me, it's my pleasure to answer your questions.

Sanyam Bhutani: Today, you're working at Google AI Research and you've created one of the most popular libraries in Deep Learning.

Could you tell us how you got started? What got you interested in Deep Learning?

François Chollet: I've been into AI in general for a long time, but I got interested in the specific problem of learning stacked, modular, hierarchical representations for visual perception in the late 2000s as a student. Back then, Jeff Hawkins had written a pretty thought-provoking book (On Intelligence) about, among other things, hierarchical information processing in the brain, and some folks at MIT had been working on hierarchical models for visual perception inspired by the human visual cortex, the HMAX family of models.

These ideas struck me as profound and correct, so I started to work on my own algorithms. I wasn't using neural networks back then, I started with stacked feature learners based on matrix factorization. I wanted to learn not just hierarchies of visual features, but hierarchies of visual-temporal features, incorporating correlations over time, not just space. I had some initial practical successes in mid-2012, while I was doing research at the University of Tokyo—I applied my setup to unsupervised few-shot recognition of hand gestures and got some really nice results. A few months later, Hinton's lab won the ImageNet competition (which my lab at the University of Tokyo had also entered) with a deep convnet trained on GPUs, so I got intrigued by that approach.

This is something that people have lost sight of nowadays, but end-to-end differentiable models trained with backpropagation are just one solution to the problem of learning modular-hierarchical representations for perception, and there are alternative avenues that haven't been explored very much. And that problem itself is just one of the many subproblems that compose the field of AI.

Sanyam Bhutani: Could you share what was the original motivation behind creating Keras? Did you expect it to grow so big?

François Chollet: No, I definitely didn't expect it to grow so big. I kind of expected it to make a bit of a splash in the small community of people using deep [learning](#) at the time, in March 2015 (a few thousands of people), but back then no one really expected deep learning would gather such interest over the following few years.

I started Keras for my own use, pretty much. I was doing natural language processing research at the time, in my free time, and I was looking for a good [tool](#) to work with RNNs. The LSTM algorithm was still largely off the radar

back then—now it's everywhere—but as neural networks were gaining prominence in the machine learning world, a number of people started investigating the use of LSTM for natural language processing. There was no reusable open-source implementation of LSTM at the time, as far as I can tell. So I made one, using Theano, which I had been using for the previous year. And then I made more layers. And that turned into a framework. I named it Keras, I open-sourced it, and it grew from there.

What made it different at the time was this: it was pretty accessible and easy to use compared to other options, it supported *both* RNNs and convnets (that was a first, I believe), and it made you define models via Python code rather than via configuration files (which had been the most popular approach previously, in particular, that was the approach of Caffe and PyLearn2).

Sanyam Bhutani: What are your thoughts about Keras being made the default API in TensorFlow 2.0. Why do you feel it is necessary?

François Chollet: TensorFlow is an extremely powerful framework, but it has long suffered from usability issues, in particular a sprawling and sometimes confusing API. TensorFlow 2 fixes these issues in a big way. Two things are at the center of this effort: eager execution, and the Keras API. Eager execution brings an imperative coding style to TensorFlow, making it more intuitive and easier to debug. The Keras API consolidates usage patterns into one coherent spectrum of really productive and enjoyable workflows, suited to a variety of user profiles, from research to applications development to deployment. I'm really excited about what we're about to release. But you'll see for yourself soon!

Sanyam Bhutani: Outside of TF and Keras, what other frameworks do you think look promising?

François Chollet: I think MXNet looks very promising, together with its high-level API, Gluon, which is very inspired by Keras and Chainer. MXNet leverages a lot of the same ideas as TensorFlow 2—a blend of eager and symbolic execution. Together with TensorFlow, it's one of the few frameworks that are actually production-grade and scalable. And MXNet has a lot of engineering firepower behind it—there's a large team at Amazon working on it. It's a serious project with some very good ideas and solid execution.

Sanyam Bhutani: For the readers and the beginners who are interested in working on Deep Learning with the dream of working at Google someday, what would be your best advice?

François Chollet: I don't think you should tie your dreams to external markers of status, like working for a specific big-name company, or making a specific amount of money, or attaining a specific job title. Figure out what you value in life, and then stay true to your values. You'll never have to regret a single decision.

Sanyam Bhutani: Could you tell us what does a day in your life looks like?

François Chollet: It's not very glamorous, mostly reviewing a lot of code, talking to people, writing design docs, shipping code. I still write quite significant amounts of code.

Sanyam Bhutani: It's a common belief that you need major resources to produce significant results in Deep Learning.

Do you think a person who does not have the resources that someone at Google might have access to, could produce significant contributions to the field?

François Chollet: There are some category of problems that require industry-scale training resources, for sure. But there is no shortage of problems where a single GPU is all you need to make significant progress. The main thing that's holding back AI research right now is not a lack of hardware, it's a lack of diversity of thought. If you have limited resources, don't spend your time worrying about GPUs, rather, spend it worrying whether you're working on the right problem and asking the right questions.

Sanyam Bhutani: You've been an advocate of Ethics in "AI". Could you share a few areas that we must pay attention to when building "AI Products"?

François Chollet: Other people have talked about the ethical issues of machine learning far better than I could. Kate Crawford, for instance, or Meredith Whittaker. I think everyone interested in this should check out their work.

Sanyam Bhutani: Do you feel Machine Learning has been overhyped?

François Chollet: Definitely, to some extent. I think that machine learning is, in a way, simultaneously overhyped and underrated. On one hand, people tend to vastly overestimate the intelligence and the generalization power of current machine learning systems, perceiving machine learning as a kind of magic wand that you can wave at arbitrary problems to make them disappear. This is, of course, largely false, there is very little actual intelligence in our algorithms, and their scope of application is extremely narrow. But at the same time, most people still underestimate how much can be achieved with the relatively crude systems we have today, if we apply them systematically to every problem they can potentially solve. Machine learning is, in a way, the steam power of our era: a pretty basic mechanism that nonetheless has the potential to profoundly change the world when used at scale.

Sanyam Bhutani: Do you feel a Ph.D. or Masters level of expertise is necessary or one can contribute to the field of Deep Learning without being an “expert”?

François Chollet: Plenty of significant contributors in the field of deep learning today don't have a Ph.D. To contribute meaningfully to a field, whether with systems development or with novel research, you absolutely do need to have a certain level of expertise. But you can gain expertise without going through a Ph.D. program, obviously, and having a Ph.D. is not actually a guarantee that you've developed meaningful expertise in anything—in theory it should, but as far as I can tell, reality doesn't align very well with that theory. In fact, unless you aim at becoming an academic, I don't think getting a Ph.D. is the best path to gaining expertise. The best path is a path that gets you to grow fast, open-endedly. And you will learn the fastest by working on a large variety of projects in situations of teamwork and close mentorship from experts. In practice, the typical Ph.D. program looks nothing like that.

Sanyam Bhutani: Before we conclude, any advice for the beginners who feel overwhelmed to even get started with Deep Learning?

François Chollet: In 10 years you'll be able to buy a textbook that will neatly sum up every AI advance that has happened from 2010 to 2020. The flood of content being published today may look important, but most of it is noise. Focus on the big questions.