



Carlos E. Perez

Deep Learning Patterns, Methodology and Strategy @ IntuitionMachine.com

Jan 13 · 8 min read

The Many Tribes of Artificial Intelligence



Credit: <https://unsplash.com/search/tribal?photo=UTKbCY454pM>

One of the biggest confusions about “Artificial Intelligence” is that it is a very vague term. That’s because Artificial Intelligence or AI is a term that was coined way back in 1955 with extreme hubris:

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire.

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions, and concepts, solve kinds of problems now reserved for humans, and improve themselves.

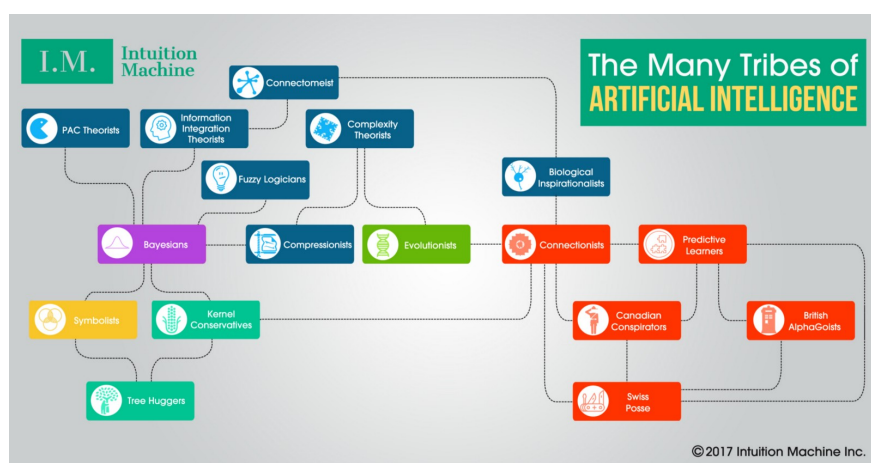
- Dartmouth AI Project Proposal; J. McCarthy et al.; Aug. 31, 1955.

AI is over half a century old and carries with it too much baggage. For a very long time, AI was dominated by Symbolists, that is rule-based systems that had “Zero Learning”. In the 1980’s a new kind of AI began to emerge, we termed this Machine Learning. Finally, we had at

least “Simple Learning”. The big disruption, however, occurred this decade, when we stumbled upon “Deep Learning” and ever since it has been taking no prisoners.

This is, of course, a grossly simplified history of AI. There are actually many different approaches or tribes in AI. Pedro Domingo’s in his book, the Master Algorithm, talks about five different “tribes”. Not to be outdone, A YCombinator user “solidrocketfuel” posts about at least “21 different cultures”.

It is important for anyone that plans on doing AI to understand that there are differences in the approaches of the different tribes of AI. AI is not a homogenous field, but rather a field in constant tribal warfare. Here’s an overview:



And a quick breakdown:

Symbolists—Folks who used symbolic rule-based systems to make inferences. Most of AI has revolved around this approach. The approaches that used Lisp and Prolog are in this group, as well as the SemanticWeb, RDF, and OWL. One of the most ambitious attempts at this is Doug Lenat’s Cyc that he started back in the 80’s, where he has attempted to encode in logic rules all that we understand about this world. The major flaw is the brittleness of this approach, one always seems to find edge cases where one’s rigid knowledge base doesn’t seem to apply. Reality just seems to have this kind of fuzziness and uncertainty that is inescapable. It is like playing an endless game of Whack-a-mole.

Evolutionists- Folks who apply evolutionary processes like crossover and mutation to arrive at emergent intelligent behavior. This approach is typically known as Genetic Algorithms. We do see GA

techniques used in replacement of a gradient descent approach in Deep Learning, so it's not a approach that lives in isolation. Folks in this tribe also study cellular automata such as Conway's Game of Life [CON] and Complex Adaptive Systems (CAS).

Bayesians—Folks who use probabilistic rules and their dependencies to make inferences. Probabilistic Graph Models (PGMs) are a generalization of this approach and the primary computational mechanism is the Monte-Carlo method for sampling distributions. The approach has some similarity with the Symbolist approach in that there is a way to arrive at an explanation of the results. One other advantage of this approach is that there is a measure of uncertainty that can be expressed in the results. Edward is one library that mixes this approach with Deep Learning.

Kernel Conservatives—One of the most successful methods prior to the dominance of Deep Learning was SVM. Yann LeCun calls this glorified template matching. There is what is called a kernel trick that makes an otherwise non-linear separation problem into one that is linear. Practitioners in this field live in delight over the mathematical elegance of their approach. They believe the Deep Learners are nothing but alchemists conjuring up spells without the vaguest of understanding of the consequences.

Tree Huggers—Folks who use tree-based models such as Random Forests and Gradient Boosted Decision Trees. These are essentially a tree of logic rules that slice up the domain recursively to build a classifier. This approach has actually been pretty effective in many Kaggle competitions. Microsoft has an approach that melds the tree based models with Deep Learning.

Connectionists—Folks who believe that intelligent behavior arises from simple mechanisms that are highly interconnected. The first manifestation of this were Perceptrons back in 1959. This approach died and resurrected a few times since then. The latest incarnation is Deep Learning.

There are many sub-approaches under Deep Learning. This includes:

The Canadian Conspirators—Hinton, LeCun, Bengio et al. End-to-end deep learning without manual feature engineering.

Swiss Posse—Basically LSTM and that consciousness has been solved by two cooperating RNNs. This posse will have you lynched if you

ever claim that you invented something before they did. GANs, the “coolest thing in the last 20 years” according to LeCun are also claimed to be invented by the posse.

British AlphaGoist—Conjecture that AI = Deep Learning + Reinforcement Learning, despite LeCun’s claim that it is just the cherry on the cake. DeepMind is one of the major proponents in this area.

Predictive Learners—I’m using the term Yann LeCun conjured up to describe unsupervised learning. The cake of AI or the dark matter of AI. This is a major unsolved area of AI. I, however, tend to believe that the solution is in “Meta-Learning”.

In addition to the above mainstream approaches, we also have:

Compressionists—Cognition and learning are compression (Actually an idea that is shared by other tribes). The origins of Information theory derives from an argument about compression. This is a universal concept that it is more powerful than the all too often abused tool of aggregate statistics.

Complexity Theorists- Employ methods coming from physics, energy-based models, complexity theory, chaos theory and statistical mechanics. Swarm AI likely fits into this category. If there’s any group that has a chance at coming up with a good explanation why Deep Learning works, then it is likely this group.

Fuzzy Logicians—This approach was once quite popular, but for some reason, I haven’t heard much about it as of late. One would think that there would be a little more interest here considering the success of the also ‘fuzzy’ approach of Deep Learning. There was a recently published result that showed the use of Fuzzy rules defeating a fighter pilot in a mock dogfight.

Biological Inspirationalists—Folks who create models that are closer to what neurons appear in biology. Examples are the Numenta folks and the Spike-and-Integrate folks like IBM’s TrueNorth chip.

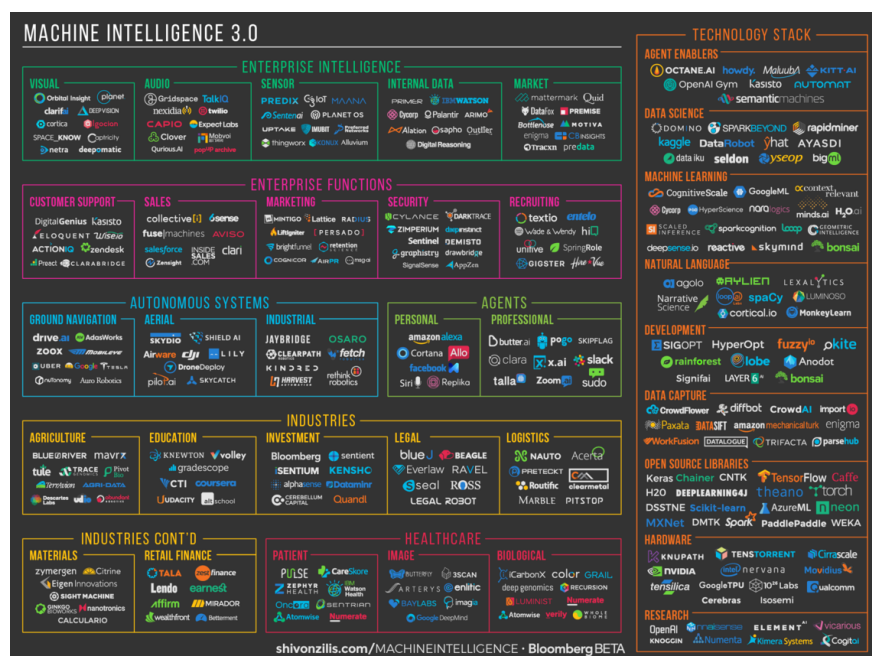
Connectomeist—Folks who believe that the interconnection of the brain (i.e. Connectome) is where intelligence comes from. There’s a project that is trying to replicate a virtual worm and there is some ambitious heavily funded research [HCP] that is trying to map the brain in this way.

Information Integration Theorists—Argue that consciousness emerges from some internal imagination of machines that mirrors the causality of reality. The motivation of this group is that if we are ever to understand consciousness then we have to at least start thinking about it! I, however, can't see the relationship of learning and consciousness in their approach. It is possible that they aren't related at all! That's maybe why we need sleep.

PAC Theorists—Are folks that don't really want to discuss Artificial Intelligence, rather prefer just studying intelligence because at least they know it exists! Their whole idea is that adaptive systems perform computation expediently such that they are all probably approximately correct. In short, intelligence does not have the luxury of massive computation.

In summary, there really is a bewildering array of alternative approaches to AI. I am certain that there are other approaches that I have missed. Some approaches are in opposition to each other, while other can be used together synergistically. However, what I want to point out is that a bit of an understanding of what is out there can help you navigate this space.

If we do take a look at Shivon Zillis' Machine Intelligence Landscape:



Credit: <https://www.oreilly.com/ideas/the-current-state-of-machine-intelligence-3-0>

There are a ton of companies that all claim to be doing AI. You have to ask a more pointed question. What sort of AI is each of these firms

using? Because the stark reality here is that not all AI are the same. Said differently, “Some AI are more equal than other AI”. Our opinion is that Deep Learning related approaches have a disproportionate monopoly on the upside. The simple reason is: “It is the learning, stupid!” If your AI approach does not have a strong mechanism for learning, then you will forever be doomed to Doug Lenat’s fate. That is, having to write all the rules by hand (for 30 years)! The other approaches tend to be pretty much dead end approaches. It is critical that the AI approach has a way to learn or alternative mechanically bootstrap internal rules.

One of the most effective approaches has been to use Deep Learning in combination with other algorithms. We have seen this in the AlphaGo implementation that used a Monte-Carlo Tree Search technique in combination with Deep Learning. The integration of a symbolic approach with Deep Learning is also very promising considering that they have complementary strengths and weaknesses. Looking forward into the future, it’s all going to be Deep Learning, one AI to rule them all. However, more likely, Deep Learning combined with some other AI approach will be just as promising. Ignoring this reality is a surefire way to ensure one’s own extinction.

An update from Davos. In an interview [WEF] with Sergey Brin of Google, Brin remarks that despite being a few paper clips away from Google Brain’s Jeff Dean, he did not see Deep Learning coming. If Brin himself who created Google could not have predicted the advance of Deep Learning, then what can we say about the many other practitioners from other competing AI tribes? That is people with a vested interest? To conclude, do not trust your resident AI expert, he just might belong to a different tribe.



For more on this, read “The Deep Learning Playbook”

I found this: [A Review of 40 Years of Cognitive Architecture Research: Focus on Perception, Attention, Learning and Applications.](#)

Update: Wired magazine has an articles about the “[AI Factions get Feisty.](#)”

[CON]<https://niginsblog.wordpress.com/2016/03/07/new-spaceship-speed-in-conways-game-of-life/>

[HCP]<https://www.humanconnectome.org/>

[WEF]<http://www.weforum.org/agenda/2017/01/google-sergey-brin-i-didn-t-see-ai-coming/>