

## Some Future Directions

*Geoffrey Fox 13 March 2020*

We expect that a major driver over the next 30 years will be deep learning which will continue to replace other machine learning algorithms as the best practice approach to AI. Further, it will be applied to many more scenarios. Of course, deep learning itself will continue to improve, probably in major ways but this will only strengthen our conclusion. An important aspect of deep learning is that one does not need to specify the model in detail. Rather the array of neurons and weights learns the model. This feature will continue and increase in power. Now one chooses a general model structure and then optimizes weights with training while at the same time using hyperparameter searches to explore details of the network. As time progresses we get more and more data as well as improving our understanding of the model building. This will allow us to combine hyper-parameter and parameter optimization into a single activity (advancing today's AutoML) which will include a search over different model structures so it will no longer be necessary to have an expert decide on the choice of for example recurrent versus fully connected versus convolutional versus graph neural network at each network stage. The huge increases in data and trainable parameters will lead to dramatic increases in required computational resources and it is likely that this will be delivered by specialized AI hardware; what we can get from Moore's law is insufficient. This scenario where deep learning (or whatever its successor is called) can learn everything trainable will lead to new computer architectures where all activities from network transmission to operating system to big data analytics to yottascale simulations will be bathed in a pervasive deep learning aether. This will allow us to optimize "everything" and dramatically improve knowledge discovery. The aether is distributed stretching across data centers and the edge and needs a new operating environment, which will be a major middleware research area. Algorithms will also be very active as although we will be dealing with flexible self-learning systems, the aether will perform more effectively and in a more explainable way if it is designed in a clever way.

How do we advance this vision? An urgent need is to increase research in both the application of and core technology for deep learning which is currently used much less in science than in industry. We also need to begin the Aether operating system with more advanced parameter and hyper-parameter training environments. Probably we should move from the current ad-hoc and genetic algorithm searches to ones with their own custom deep learning agents. This should be pursued with much greater collaboration between science and industry. Further, we need to review the needed education and training which itself will benefit from deep learning. I expect the academic field of data science to gain in importance.