December, 2017

The Opinion of Machines

Curtis E.A. Karnow

Preprint. Final version scheduled for XIX COLUMBIA SCIENCE & TECHNOLOGY LAW REVIEW (Dec. 2017)

Abstract. The note reviews the application and operation of artificial neural networks, and proposes a test for the admissibility of opinions generated by such artificial intelligence systems. The discussions includes a review of the admissibility of computer stored and computer generated evidence, including computer simulations.

The Opinion of Machines

By

Curtis E.A. Karnow[*]

> *People understand the linear algebra behind deep learning [neural networks]. But the models it produces are less human-readable. They're machine-readable. They can retrieve very accurate results, but we can't always explain, on an individual basis, what led them to those accurate results*.[1]

> *When I watch these games, I can't tell you how tense it is. I really don't know what is going to happen*.[2]

Software performs many functions far better than humans. A specific software architecture, neural networks, not only takes advantage of the virtually perfect recollection and much faster processing speeds of any software, but also teaches itself and attains skills no human could program directly. We rely on these neural nets for medical diagnoses, financial decisions, weather forecasting, and many other crucial real-world tasks. In 2016, a program named AlphaGo beat the top rated human player of the game of Go.[3] Only a few years ago, this had been considered impossible.[4] High level Go requires remarkable

---

[*] Judge of the Superior Court, County of San Francisco.

[1] Chris Nicholson, "AI is transforming Google Search -- The rest of the Web is next," *Wired* (February 4, 2016) <https://www.seroundtable.com/google-dont-understand-rankbrain-21744.html>

[2] David Silver, one of AlphaGo's creators. Cade Metz, "What The Ai Behind Alphago Can Teach Us About Being Human," *Wired* (May 19, 2016) < https://www.wired.com/2016/05/google-alpha-go-ai/>. Silver is the author of e.g. sources cited below at notes 3, 134. "There are a lot of things I don't understand," Nick Sibicky, Lecture 256, on the games AlphaGo plays against itself. Sibicky is a strong Go player, and a professional Go instructor. <https://www.youtube.com/watch?v=yfUzW0gH8ts?>.

[3] David Silver et al., "Mastering the game of Go with deep neural networks and tree search," 529 NATURE 484 (28 January 2016) <http://web.iitd.ac.in/~sumeet/Silver16.pdf>.

[4] E.g., Alan Levinovitz, "The Mystery Of Go, The Ancient Game That Computers Still Can't Win," *Wired* (May 12, 2014) ("But the fact is that of all the world's deterministic perfect information games – tic-tac-toe, chess, checkers, Othello, xiangqi, shogi – Go is the only one in which computers don't stand a chance against humans."). See George Johnson, "To Test a Powerful Computer, Play an Ancient Game," *The New York Times* (July 29, 1997)

skills not just of calculation, which computers obviously excel at, but more critically judgment and intuition, pattern recognition, the weighing of ineffable considerations such as balance.[5] These cannot be directly programmed. But AlphaGo's neural network[6] trained itself with many thousands of games (and later, millions)—far more than any individual human could ever play[7]—and now routinely beats all human challengers.[8] Because it learns, and concomitantly modifies itself in response to experience, such a network is termed *adaptive*.[9]

As detailed below, neural networks are used throughout industry and science. They are proposed for missile launch and interception.[10] If these systems are so useful, presumably juries could rely on their expert opinions as well. Enabling the admission of what we might call machine opinion evidence requires a review of the requirements of expert opinion and a trial judge with the ability to rule on admissibility and ensure the opinion is correctly framed for the jury. There are two aspects to this ability: judges must have enough knowledge to handle the technical issues, and must have the authority to decide that the software is scientifically reliable, with an appreciation of the risks. Most judges probably do not have this knowledge, and current California law may not countenance that sort of admissibility analysis. This note may assist on those two problems.

Judges and lawyers well know the apparent authority of experts--those professing independence and authority-- used to sway juries. There are high risks that juries will invest computer systems with even greater authority as free of bias, independent of the parties, and error-free.[11] Most especially in this

---

http://www.nytimes.com/1997/07/29/science/to-test-a-powerful-computer-play-an-ancient-game.html ("'It may be a hundred years before a computer beats humans at Go -- maybe even longer,' said Dr. Piet Hut, an astrophysicist at the Institute for Advanced Study in Princeton, N.J.").

[5] This is so because the number of possible permutations is practically infinite (the number of possible Go games far, far exceeds the number of atoms in the universe, < http://senseis.xmp.net/?NumberOfPossibleGoGames>), and mere calculation cannot beat even a modestly good human player. This is as opposed to chess, which has far fewer options than Go. For chess, a so-called brute force approach can beat top human players. IBM, Frequently Asked Questions: Deep Blue <https://www.research.ibm.com/deepblue/meet/html/d.3.3a.shtml>. See Cade Metz, "In A Huge Breakthrough, Google's Ai Beats A Top Player At The Game Of Go," *Wired* (January 27, 2016) ("When Deep Blue topped world chess champion Gary Kasparov in 1997, it did so with what's called brute force. In essence, IBM's supercomputer analyzed the outcome of every possible move, looking further ahead than any human possibly could. That's simply not possible with Go.") <https://www.wired.com/2016/01/in-a-huge-breakthrough-googles-ai-beats-a-top-player-at-the-game-of-go/>. The 1997 New York Times article quoted in note 4 provides a good explanation of the differing complexities as between Go and chess.

[6] For a general discussion of AlphaGo's neural net see <https://www.tastehit.com/blog/google-deepmind-alphago-how-it-works/>. Neural networks are so called because they operate in layers, each with different function. Ian Goodfellow, et al., DEEP LEARNING 6 (2016).

[7] See generally, Cade Metz, "What The Ai Behind Alphago Can Teach Us About Being Human," *Wired* (May 9, 2016) available at <https://www.wired.com/2016/05/google-alpha-go-ai/>.

[8] See the website for the American Go Association, reporting on January 4, 2017 that AlphaGo was confirmed as the secret player defeating 50 of the top Go players in the world. <http://www.usgo.org/>.

[9] Mohamad Hassoun, "What is a neural network and how does its operation differ from that of a digital computer?," *Scientific American* <https://www.scientificamerican.com/article/experts-neural-networks-like-brain/>

[10] E.g., Jin-ke Xiao at al., "Improved Clonal Selection Algorithm Optimizing Neural Network for Solving Terminal Anti-missile Collaborative Intercepting Assistant Decision-Making Model," 644 *Communications in Computer and Information Science* 216-231 (22 September 2016); Eric Walh, et al., "Non-linear receding horizon control based real-time guidance and control methodologies for launch vehicles," Aerospace Conference, 2016 IEEE (2016); M. B. McFarland, et al., "Adaptive nonlinear control of agile antiair missiles using neural networks," 8 *Control Systems Technology*, *IEEE Transactions* 749-756 (2000) <http://www.au.af.mil/au/awc/awcgate/cst/bh_manor.pdf/>.

[11] Erin E. Kenneally, "Gatekeeping Out of the Box: Open Source Software As A Mechanism to Assess Reliability for Digital Evidence," 6 VA. J.L. & TECH. 13 (2001) ("digital evidence may carry an aura of infallibility in the public's eyes").

context must trial judges carefully undertake their gate-keeping functions[12] and ensure only reliable evidence gets to the jury.

While there are some court opinions involving neural networks such as patent cases,[13] I have found no state or federal case discussing the admissibility of what I will term machine opinion, that is, an evidentiary statement generated by software which no human can fully explain. A number of commentators, however, have suggested the evidence be admissible. They have explored, for example, facial recognition software which reports the probability that a fuzzy picture is that of a defendant; that is, in circumstances where no human could make a similar estimate.[14] Commentators have also suggested applications to prove fraud in the health care industry, which would require pattern detections in very large amounts of data.[15] Further examples are provided below.

A. Neural networks

1. *An Introduction for Lawyers: Predictive Coding*

Many lawyers are already familiar with the basic technology, because they use neural networks in their technology assisted review (TAR) of voluminous electronic documents.[16] With productions of millions of emails and other documents, it is not only futile to have humans review these pages, but TAR searches for relevant items are usually cheaper, and almost always more accurate. The systems use so-called predictive coding. They are trained on a set of documents, selected by humans as representative of the universe of documents at issue. The training is done with a so-called 'seed set' of documents. Then the system is provided a sample of the general production, and offers up its opinion on what is relevant and what is not. Humans train the system by noting errors, and the system then iteratively refines its ability to discriminate. It does this by, in effect, weighing aspects of the documents, such as key words and series of words, to generate a probability that the item is relevant or that it is not. When the system is sufficiently accurate with respect to its training (or "control set") documents, it is then let loose on the entire corpus of the production—the many millions of documents at issue—and marks those which are in its opinion (as it were) responsive. "Predictive discovery is faster, cheaper and more accurate than traditional discovery approaches."[17]

[12] *Sargon Enterprises, Inc. v. Univ. of S. Cal*., 55 Cal. 4th 747 (2012); *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993). To be clear, this note is addressed specifically to the threshold issue of the *admissibility* of opinions. While reliability of an opinion is or should be the most important factor in tests for both admissibility (e.g., *Wendell v. GlaxoSmithKline LLC,* __F,3d__, __(No. 14-16321, 2017 WL 2381122, at *6) (9th Cir. June 2, 2017)) and subsequent acceptance by the trier of fact (the judge or the jury), admissibility is distinct from whether the opinion is ultimately treated as convincing by the trier of fact.

[13] E.g., *Neuromedical Sys., Inc. v. Neopath, Inc.,* No. 96 CIV. 5245 (JFK), 1998 WL 264845, at *4 (S.D.N.Y. May 26, 1998).

[14] E.g., John Nawara, "Machine Learning: Face Recognition Technology Evidence in Criminal Trials," 49 U. LOUISVILLE L. REV. 601 (2011). There are interesting Confrontation Clause issues. Joseph Clarke Celentino, "Face-to-Face with Facial Recognition Evidence: Admissibility Under the Post-Crawford Confrontation Clause," 114 MICH. L. REV. 1317 (2016).

[15] Neil Issar, "More Data Mining for Medical Misrepresentation? Admissibility of Statistical Proof Derived from Predictive Methods of Detecting Medical Reimbursement Fraud," 42 N. KY. L. REV. 341 (2015). Other suggestions are found in e.g., Andrea Roth, "Machine Testimony," 126 YALE L.J. 1972, 2021 (2017).

[16] See e.g., Shannon Brown, "Peeking Inside the Black Box: A Preliminary Survey of Technology Assisted Review (Tar) and Predictive Coding Algorithms for Ediscovery," 21 SUFFOLK J. TRIAL & APP. ADVOC. 221 (2016); Aaron T. Goodman, "Predictive Coding and Electronically Stored Information Computer Analytics Combat Data Overload," ARIZ. ATT'Y (July/August 2016).

[17] Joseph H. Looby, "E-discovery – Taking Predictive Coding Out of the Black Box," FTI JOURNAL, <http://ftijournal.com/article/taking-predictive-coding-out-of-the-black-box-deleted>, relying on "Technology-Assisted Review in E-discovery Can Be More Effective and More Efficient than Exhaustive Manual Review," by

I make these observations: (1) No one knows why the system selects a document; that is, once the system is trained, there is no script that can be provided to a human sorter to imitate the system's selection of documents. There is no way to accurately summarize the criteria used. (2) Nevertheless, parties rely on predictive coding in very high stakes litigation. It is reliable.

2. *Under the Hood: Hidden Layers*

Having noted the legal profession's general familiarity with a sort of neural network, I provide a short introduction to a typical mechanism of these programs.

At the risk of conflating uses of the term 'expert,' I contrast neural nets with a classic "expert system". The classic expert system is simply a collection of rules, expressly preprogrammed by a human. For example we can imagine a car repair (or medical) expert system which asks a series of scripted questions, and then spits out an answer. A human expert scripted the questions and created the matrix such that a certain set of responses generate a scripted output.[18] We might say the operations are 'hard coded' into the software.[19] Some legal work can probably be done with these systems.[20] A human can fully explain the operation of this sort of expert system.

Simple machine learning systems learn from data presented from the real world. These normally require structured data, which means humans must in effect interpret the data from the world, rendering it into representations acceptable to the program.[21] Representational learning systems, and in particular deep learning systems, do not require this human intervention. These programs can be exposed to data from the real world and be taught—and later, teach themselves—the relationship between (i) raw data and (ii) higher level representations and abstract concepts.[22] Neural networks are a type of representational learning system; some of them are deep, and some are shallow, as described below. These solve problems that cannot be solved by fixed programs written by humans.[23]

Neural networks are arranged such that the actual operation, the weighing of probabilities, is not perceived by humans. Humans do not fix the way in which elements are weighed, and they usually do not even identify *which* elements are weighed. The nets organize themselves. (As we will see, recent results are even more surprising: networks have trained themselves on unlabeled data to recognize for

---

Maura R. Grossman and Gordon V. Cormack, 17 RICHMOND JOURNAL OF LAW AND TECHNOLOGY 11 (2011). *See also,* e.g. Veritas, "Predictive Coding Defensibility" (2015) ("Despite the widespread misconception that linear review is the electronic discovery process "gold standard," exhaustive manual review is surprisingly inaccurate, considering its high cost. Academic research on legal review as part of the TREC Legal Track has shown linear review is often only 40-60 percent accurate. Predictive coding technology involves an iterative process that senior attorneys follow to train software on review criteria, creating a mathematical model that predictive coding software uses to generate "predictions" of how the remaining documents would otherwise be tagged if reviewed by an experienced attorney. Studies show that predictive coding can achieve much higher levels of accuracy at a fraction of the time and cost.") <https://www.veritas.com/content/dam/Veritas/docs/white-papers/21290290_GA_ENT_WP-Predictive-Coding-Defensibility-Measuring-Accuracy-with-Random-Sampling-EN.pdf>.

[18] Frederick P. Brooks, THE MYTHICAL MAN-MONTH 191 (1995). Brooks in his classic text (originally published in 1975) calls these now relatively simple expert systems 'inference engines'. While the same term could be used for neural networks, the means of inference, and their flexibility, differ profoundly.

[19] DEEP LEARNING at 2.

[20] McCarty, L. Thorne, "Reflections on Taxman: An Experiment in Artificial Intelligence and Legal Reasoning," HARV.L.REV. 90 (1997).

[21] DEEP LEARNING at 2-3.

[22] Id. at 4-5.

[23] Id. at 96.

example faces and cats—that is, the systems make these discriminations without first fed examples of the items to be discriminated.[24] More on this below.) These systems use statistics and algorithms derived from probability theory[25] to navigate uncertain and ambiguous data to generate results, and then teach themselves in effect to revise their own algorithms in order to increase accuracy.

A good example of a neural network is that used to do image analysis, such as recognizing faces, a bus, or other features in pictures.[26] The system first accepts input. In our example, this is a series of pixels which for simplicity's sake will be either black or white, on or off. We might have then a grid, perhaps 200 by 200 (i.e. 4,000) pixels or dots. The first task is then to recognize the input are on or off—let's call that the work of the first layer of nodes. From this we may move to the second task: whether there are edges. Three black dots in a row might be an edge; perhaps seven are very likely to be an edge, and 10 in a row are extremely likely to be so. Edge detection might be then the second layer of processing. Depending on how the node is adjusted (more on that below) some of the nodes might 'vote' that there is an edge; or not. We don't yet know if we have a face or a baseball.

The second layer's output—"we have edge" or, "no edge here"—is then the input for the next layer; we might call that a shape detector, or eye detector, for example, At this third layer, the edges are determined to either fit together in a certain shape, or not. The output here might be something like, "we have an eye" or nose, etc., or some other elemental shape. That output is the input of the next (fourth) layer, which we might call a face recognition layer, which, given the input of eyes and noses etc.--or some of them (and we'll get back to this shortly)—creates a final output: "We have face" or "no face here" or, if the penultimate layer were trained to look for things like wheels, side panels, cabs, and so on, it might report "it's a truck." At each layer, the input is likely to vary greatly: edges come in all sorts of shapes and sizes, sometimes they are manifest in a few pixels and sometimes many more; these edges at subsequent layers to a greater and lesser extent conform to an eye, or nose, or wheel, or head, and so on, and those elements in turn conform to a truck, or baseball, to greater or lesser extent. The output of a layer to the next layer is probabilistic; it might take only a weak probability to send a "yes" up the chain so to speak, or it might take a high degree of certainty to send that "yes, it's an edge" or "yes, this is wheel." A layer may have some but not all of the input it needs to be certain of a conclusion, and so in effect its nodes vote on the *degree* of certainly about the conclusion. The nodes in the network that, in the end, either do or do not send on a "yes" to the next layer are adjustable—and here is where the training comes in.

During training, the system makes adjustments to the nodes, assigning more or less weight to input from earlier layers. In the classic training session, the system is fed a large number of labeled pictures (or for predictive coding in the ESI context, documents), and is provided human feedback. It is told if it got the decision right, or not. If not, the system experiments internally adjusting the weights of its nodes until it maximizes the number of right estimates or final outputs. The classic example is a 'back propagational neural network' where the final output error is used iteratively to go 'back' and tweak the nodes' weightings, run another effort, and note the extent to which the output improves. Whether technically correct or not, the comparisons to human learning are obvious:[27] we say children are taught that various

---

[24] Quoc V. Lee, et al., "Building High-Level Features Using Large Scale Unsupervised Learning," Proceedings of the 29th International Conference on Machine Learning, Edinburgh, Scotland, UK (2012), <https://static.googleusercontent.com/media/research.google.com/en//archive/unsupervised_icml2012.pdf>; <https://arxiv.org/pdf/1112.6209.pdf>.

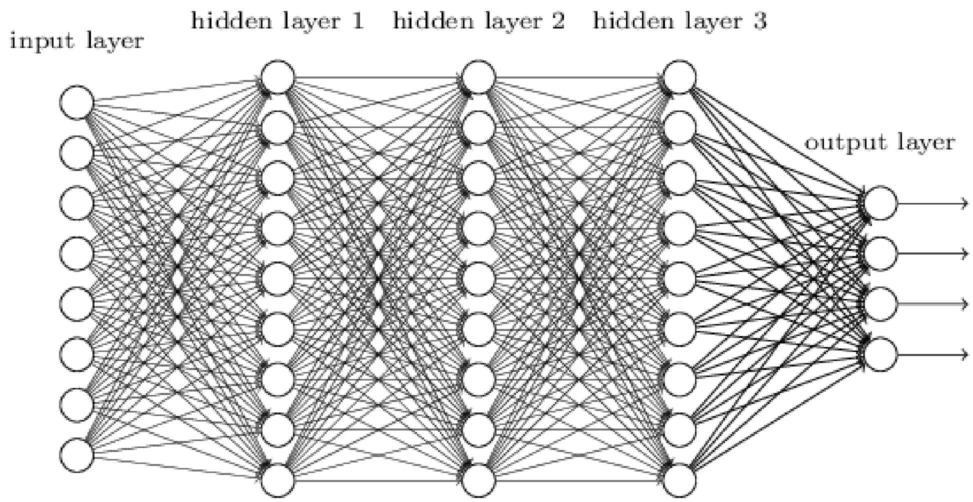
[25] DEEP LEARNING at 52 *et seq*.

[26] See generally, DEEP LEARNING at 6, figure 1.2.

[27] Our intuition that artificial neural networks mimic our biological ones may be right. David Hubel and Torstein Wiesel were awarded the 1981 Nobel Prize in physiology or medicine for work on the information processing systems in the visual cortex which uses the equivalent of hidden layers of neural network. <http://www.nobelprize.org/nobel_prizes/medicine/laureates/1981/press.html>. Others caution that the brain is like a neural network only by way of analogy and metaphor. E.g., Chris Chatham, "10 Important Differences Between

things are dogs, or cats, by repeatedly correcting the child's output statements 'doggie!' or 'kittie!' until, by and large, the output is correct. As with neural nets, we can measure, and ultimately have some faith, in the accuracy of the output, but we have no idea what the internal state of the network (or child's brain) looks like, or exactly why it is so. In a network, these are just a very large number of weights, i.e. numbers. These layers in between the initial input and the final output are thus often referred to as "hidden layers".[28] As the system traverses the layers from the raw data input to the final output, it reaches conclusions about increasingly complex and abstract concepts.[29]

Here's a simple diagram of a five layer network:[30]



Supervised networks train using labeled data, and then estimate answers from new input. As I have noted, some neural networks can train themselves, taking great advantage of the amount of digitized data which has vastly increased in recent years.[31] 'Big data' allows programs much room to train and self-correct their mechanisms. While the line between supervised and unsupervised learning is not fixed,[32] unsupervised learning examines the unlabeled data, compares it to random data, and extracts a series of features which are common to the non-random data. These features are of course abstractions from the input layer. The parameters of that layer may then be fixed, and then its output examined by the next layer which treats it as input, and also extracts common features for the next level of abstraction. A simple example is a clustering program, which reviews a large amount of input, makes conclusions concerning common features, and then sorts the inputs into different groups. This can all be done with unlabeled data, and human corrective input is not required. Engineers at Google had an early incarnation of AlphaGo teach itself to recognize cats without telling it anything about cats, but by just letting it examine 13,026

Brains and Computers," *ScienceBlogs* (March 27, 2007) <http://scienceblogs.com/developingintelligence/2007/03/27/why-the-brain-is-not-like-a-co/>. The issue doesn't matter here. For those interested, artificial neural networks probably won't have the same number of neuron equivalents as humans until around 2050, DEEP LEARNING at 21; but around then artificial networks may advance very, very rapidly, unconstrained by the relatively slow processing speeds and limited storage abilities of humans.

[28] Currently, networks with about 10 layers are termed 'deep' or 'very deep'. Juergen Schmidhuber, "Deep Learning in Neural Networks: An Overview," 61 *Neural Networks* 85–117 (2015), abstract at <http://people.idsia.ch/~juergen/deep-learning-overview.html>

[29] DEEP LEARNING at 8.

[30] Michael Nielsen, NEURAL NETWORKS AND DEEP LEARNING, Ch.6 (May 2017) <http://neuralnetworksanddeeplearning.com/chap6.html>.

[31] DEEP LEARNING at 12, 19; I discuss this below § (C) (4).

[32] Id. at 100.

pictures of cats and 23,974 pictures without cats—without indicating which was which. The system eventually, on its own as it were, detected the common cat features and reported its discovery of that common entity.[33]

Unsupervised learning then improves itself as follows: assume initial unsupervised learning has created a series of higher level abstractions. For example, a system might distinguish—i.e. separately cluster, as described just above—digits from things which are not digits; or cats from things which are not cats. Now the higher layers (those generating conclusions such as "this is a digit, this is a cat") perform a top down pass in effect instructing the lower layers on what, more specifically, to look for as they make their determinations. For example, a top down pass might in effect say, 'look for about two to three strokes for digits; look for whiskers and a certain shape of ear for cats' which then iteratively improves the performance of the system as a whole.[34] The system teaches itself.

3. *Uses of neural networks*

For reason which will be express below, it is important to note the extensive use of neural networks in the physical world. We are confronted with massive amounts of data—familiarly known as 'big data'—and traditional means of analysis fail us. We want to extract patterns, and find needles in haystacks. So neural networks are used for automated bank loan application approval and credit card fraud detection, and a wide spectrum of other uses in the financial markets. They are used for medical diagnosis such as x-ray interpretation; process controls in factories, in scientific research and of course in data mining in many contexts. A more complete list is provided in this footnote.[35] One text notes these uses:

---

[33] See Quoc V. Lee, et al., at note 24 ("Contrary to what appears to be a widely-held intuition, our experimental results reveal that it is possible to train a face detector without having to label images as containing a face or not. … We also find that the same network is sensitive to other high-level concepts such as cat faces and human bodies").

[34] For a more technical discussion, but still somewhat approachable, see e.g., Geoffrey E. Hinton, "Learning multiple layers of representation," 11 *TRENDS in Cognitive Sciences* 428 (2007) <http://www.cs.toronto.edu/~hinton/absps/tics.pdf>.

[35] Financial: Stock Market Prediction; Credit Worthiness; Credit Rating; Bankruptcy Prediction; Property Appraisal; Fraud Detection; Price Forecasts; Economic Indicator Forecasts; Medical: Medical Diagnosis; Detection and Evaluation of Medical Phenomena; Patient's Length of Stay Forecasts; Treatment Cost Estimation; Industrial: Process Control; Quality Control; Temperature and Force Prediction; Science: Pattern Recognition; Recipes and Chemical Formulation Optimization; Chemical Compound Identification; Physical System Modeling; Ecosystem Evaluation; Polymer Identification; Recognizing Genes; Botanical Classification; Signal Processing: Neural Filtering; Biological Systems Analysis; Ground Level Ozone Prognosis Odor Analysis and Identification; Educational: Teaching Neural Networks; Neural Network Research; College Application Screening; Predict Student Performance; Data Mining: Prediction Classification; Change and Deviation Detection; Knowledge Discovery; Response Modeling; Time Series Analysis; Sales and Marketing: Sales Forecasting; Targeted Marketing; Service Usage Forecasting; Retail Margins Forecasting; Operational Analysis: Retail Inventories Optimization; Scheduling Optimization; Managerial Decision Making; Cash Flow Forecasting; HR Management: Employee Selection and Hiring; Employee Retention; Staff Scheduling; Personnel Profiling; Energy: Electrical Load Forecasting; Energy Demand Forecasting; Short and Long-Term Load Estimation; Predicting Gas/Coal Index Prices; Power Control Systems; Hydro Dam Monitoring; Other: Sports Betting; Making Horse and Dog Racing Picks; Quantitative Weather Forecasting; Games Development; Optimization Problems, Routing; Agricultural Production Estimates. See http://www.alyuda.com/products/neurointelligence/neural-network-applications.htm.

A quick Google Scholar review notes these splendid scientific research titles: Hydroelectric power plant management relying on neural networks and expert system integration; Use of neural network techniques in a medical expert system; Interpretation of automated perimetry for glaucoma by neural network; Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks; Neural Networks in Bioprocessing and Chemical Engineering (book); A Bayesian regularized artificial neural network for stock market forecasting; Forecasting low voltage distribution network demand profiles using a pattern recognition based expert system; Soft-sensing estimation of plant effluent concentrations in a biological wastewater

> **Detection of medical phenomena.** A variety of health-related indices (e.g., a combination of heart rate, levels of various substances in the blood, respiration rate) can be monitored. The onset of a particular medical condition could be associated with a very complex (e.g., nonlinear and interactive) combination of changes on a subset of the variables being monitored. Neural networks have been used to recognize this predictive pattern so that the appropriate treatment can be prescribed.
>
> **Stock market prediction.** Fluctuations of stock prices and stock indices are another example of a complex, multidimensional, but in some circumstances at least partially-deterministic phenomenon. Neural networks are being used by many technical analysts to make predictions about stock prices based upon a large number of factors such as past performance of other stocks and various economic indicators.
>
> **Credit assignment.** A variety of pieces of information are usually known about an applicant for a loan. For instance, the applicant's age, education, occupation, and many other facts may be available. After training a neural network on historical data, neural network analysis can identify the most relevant characteristics and use those to classify applicants as good or bad credit risks.
>
> **Monitoring the condition of machinery.** Neural networks can be instrumental in cutting costs by bringing additional expertise to scheduling the preventive maintenance of machines. A neural network can be trained to distinguish between the sounds a machine makes when it is running normally ("false alarms") versus when it is on the verge of a problem. After this training period, the expertise of the network can be used to warn a technician of an upcoming breakdown, before it occurs and causes costly unforeseen "downtime."
>
> **Engine management.** Neural networks have been used to analyze the input of sensors from an engine. The neural network controls the various parameters within which the engine functions, in order to achieve a particular goal, such as minimizing fuel consumption.[36]

Closer to the legal realm, neural nets are developed or proposed for the e-discovery uses I note above as well as, for example, detecting gunshot residue,[37] demographic analysis of crime patterns,[38] automated detection of smuggling,[39] and other uses[40] including for legal services.[41]

---

treatment plant using an optimal neural network; Soft-sensing estimation of plant effluent concentrations in a biological wastewater treatment plant using an optimal neural network; Fault Diagnosis in Internal Combustion Engines Using Extension Neural Network; Prediction of global horizontal solar irradiance in Zimbabwe using artificial neural networks.

[36] <http://documents.software.dell.com/statistics/textbook/neural-networks>.

[37] <https://www.researchgate.net/profile/Regina_Taudte/publication/275898182_Development_of_a_UHPLC_method_for_the_Detection_of_Organic_Gunshot_Residues_using_Artificial_Neural_Networks/links/557fe56f08aeea18b7797bae.pdf>.

[38] Country crime analysis using the self-organizing map, with special regard to demographic factors <http://link.springer.com/article/10.1007/s00146-013-0441-7>.

[39] Automated detection of smuggled high-risk security threats using Deep Learning <https://arxiv.org/pdf/1609.02805.pdf>.

[40] Fraud detection using self-organizing map visualizing the user profiles <http://www.sciencedirect.com/science/article/pii/S0950705114002652>; Detecting predatory conversations in social media by deep Convolutional Neural Networks, <http://www.sciencedirect.com/science/article/pii/S1742287616300731>; Neural networks for identifying drunk persons using thermal infrared imagery, <http://www.sciencedirect.com/science/article/pii/S0379073815001681>; Michael Aikenhead, "The Uses and Abuses of Neural Networks in Law," 12 Santa Clara Computer & High Tech. L.J. 31 (1996) (legal reasoning); John Nawara, "Machine Learning: Face Recognition Technology Evidence in Criminal Trials," 49 U. Louisville L. Rev. 601 (2011) (reliability of face recognition systems); Neil Issar, "More Data Mining for Medical Misrepresentation? Admissibility of Statistical Proof Derived from Predictive Methods of Detecting Medical Reimbursement Fraud," 42 N. Ky. L. Rev. 341 (2015) (statistical detection evidence).

B. Admitting Output of Software

The use of software at trial may involve issue of hearsay and reliability. Authenticity is an aspect of reliability: a document must be authenticated because otherwise it is not reliable. Hearsay objections are pertinent to some computer output and not to others. Software is used to generate simulations and animations, two very different types of evidentiary creatures with different requirements. As we will see, the rules governing the admissibility of simulations, in particular, are useful, but insufficient, in deciding whether the output of neural networks should be admitted.

But first a brief taxonomy of computer generated evidence will be useful.[42] As Justice Simons has noted, "We must distinguish from this computer-generated data [such as data associated with credit card swipes and cell phone use], a written or electronic document prepared by a person, and then electronically stored in a computer. Electronic storage does not make the document computer-generated."[43]

1. *The Filing Cabinet*

Much of what we think of as computer generated evidence (CGE)[44] is not. It is generated by humans, who input the data into computers which act simply as storage systems such as a filing cabinet. Letters, briefs, emails, PowerPoints, much (but not all) of our spreadsheets and other accounting data, and most (but not all) photographs fit in this category. Aside from most photographs, this data are in fact collections of statements by humans, and so a hearsay objection may be made.[45] The objection can be met with a showing under for example the business records exception.[46] Websites and chat room postings too are similarly data housed in digital storage cabinets: people put the words there and people can testify as to authenticity and related issues, as they would if the data came out of a filing cabinet. Databases contain human-entered data too; which fit under this rubric. Pictures posted to internet sites generally are authenticated and admitted the same way as any other photographs, that is, someone testifies either that she took the picture or someone familiar with the scene depicted testifies the photo is accurate. Circumstantial evidence may suffice for admissibility.[47]

It is true that the act of processing electronic data into legible text or photographs involves computer processing, in effect a translation of bits into a human-readable product. But the proponent of the

---

[41] John O. McGinnis et al., "The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services," 82 FORDHAM L. REV. 3041 t.a.n. 39 (2014).
[42] See generally, Gregory P. Joseph, "A Simplified Approach to Computer-Generated Evidence and Animations," 1 SEDONA CONF. J. 55 (2000).
[43] M. Simons, CALIFORNIA EVIDENCE MANUAL § 2:2 (2015 ed.).
[44] Justice Simons (see note 43) calls it CGI, computer generated information. *CGI* is also used to designate computer enhanced film effects ('computer generated imagery'), and in the interests of avoiding ambiguity I use *CGE*.
[45] Simons, note 43 at § 2:63; Gregory P. Joseph, "A Simplified Approach to Computer-Generated Evidence and Animations," 43 N.Y.L. SCH. L. REV. 875, 878 (2000); *People v. Romeo,* 240 Cal. App. 4th 931, 945 (2015) ("information residing in a computer database is still hearsay, often multilevel hearsay.").
[46] Evid. Code § 1271.
[47] "The normal rules of admissibility apply to evidence obtained from social networking and other online sites. Authentication of a photograph on a Web site may be provided by expert testimony if there is no one qualified to authenticate it from personal observation. In addition, authentication may be provided from its contents or subject matter (People v. Valdez, 201 Cal. App. 4th 1429, 135 Cal. Rptr. 3d 628 (4th Dist. 2011) (photograph from a social-networking web page alleged to have been authored by defendant was sufficiently authenticated by its content to be admissible))." Edward A. Rucker and Mark E. Overland,)." 4 CALIFORNIA CRIMINAL PRACTICE, MOTIONS, JURY INSTRUCTIONS AND SENTENCING § 48:18.10 (4th ed. March 2017). See also: Paul W. Grimm et. al., "Authenticating Digital Evidence," 69 BAYLOR L. REV. 1, 15 (2017) (circumstantial evidence widely used to authenticate); Steven Goode, "The Admissibility of Electronic Evidence," 29 REV. LITIG. 1, 24–25 (2009) (use of circumstantial evidence to authenticate in federal court).

evidence need not explain or defend this type of processing: we assume "a computer's print function has worked properly."[48] In short, "[p]rintouts are admissible and presumed to be an accurate representation of the data in the computer."[49] That being said, even when the print function presumably works correctly, the printed data may still be hearsay, because it was inputted by humans.[50]

2. *Data Created by Internal Operations*

Computers may be fed data, and based on their programming then generate new data the accuracy of which depends on the validity of the programing. The simplest examples are spreadsheet cells containing formulas which of course simply execute what we might call a mini program written by the user, such as 'multiply cell B3 with B4 and put the result here'. If the human selected the wrong cells, or directed a multiplication when it ought to have been division, the result (i.e. "profits this year were $100") will be wrong. Drawing programs can automatically generate circles and squares, but whether these are accurate depends of course on whether the algorithm is correct. In short, software may or may not have bugs,[51] and the validity of the output of course depends on this.

Typically it is suggested that because the result of internal computer processing is not a human statement, hearsay is not implicated. For example, we have this from the venerable Witkin:

> *Distinction: Computer's Internal Operations.* A printout of the results of a computer's internal operations is not hearsay evidence at all, and thus the business records exception is inapplicable. Such a printout does not represent the output of statements placed in the computer by an out-of-court declarant. With a machine, there is no possibility of conscious misrepresentation. "[T]he true test for admissibility of a printout reflecting a computer's internal operations is not whether the printout was made in the regular course of business, but whether the computer was operating properly at the time of the printout." (People v. Hawkins (2002) 98 C.A.4th 1428, 1449, 1450, 121 C.R.2d 627 [in prosecution arising from defendant's having taken source code from computer system of former employer, trial judge did not err in admitting computer printouts showing when

---

[48] *People v. Hawkins,* 98 Cal.App.4th 1428, 1450 (2002), as quoted by *People v. Goldsmith,* 59 Cal. 4th 258, 269 (2014). Of course, as with any other presumption the other side is free to attack it. But the presumption is almost always enough to get the evidence before the trier of fact (e.g. the jury). See e.g., *People v. Martinez,* 22 Cal. 4th 106, 132, 990 P.2d 563, 581–82 (2000) (problems with print out may be subject of cross examination but typically will not bar admissibility). This is because the opponent does not have evidence that the printout is inaccurate. But if she does, however, have evidence of inaccuracy, *the presumption is no longer in effect*, and the burden returns to the proponent of the print out to establish that it is, in fact, accurate. *People v. Rekte*, 232 Cal. App. 4th 1237, 1245-46 (2015).

[49] 1 B. Witkin, CALIFORNIA EVIDENCE, Hearsay § 231 (5th ed. 2012). New federal rules of evidence, effective December 1, 2017, will make it even easier in federal court to meet basic authentication requirements for computer stored data. FRE 902 (13) & (14) <http://www.uscourts.gov/sites/default/files/supreme-court-package_0.pdf>. See Paul W. Grimm et. al., "Authenticating Digital Evidence," 69 BAYLOR L. REV. 1, 39 (2017).

[50] Printouts offered for their truth usually have to qualify under some exception to the hearsay rule, such as under the business records exception. *People v. Lugashi,* 205 Cal.App.3d 632, 638 (1988); *Aguimatang v. California State Lottery*, 234 Cal. App. 3d 769, 797 (1991).

[51] Actually, all software in general use is sufficiently complex that it has bugs. "All software contains bugs or errors in the code. Some of these bugs have security implications, granting an attacker unauthorized access to or control of a computer. These vulnerabilities are rampant in the software we all use. A piece of software as large and complex as Microsoft Windows will contain hundreds of them, maybe more." Bruce Schneier, "Why the NSA Makes Us More Vulnerable to Cyberattacks: The Lessons of WannaCry," *Foreign Affairs,* SNAPSHOT May 30, 2017 <https://www.foreignaffairs.com/articles/2017-05-30/why-nsa-makes-us-more-vulnerable-cyberattacks>. The complexity of software is essential, not an accident. Frederick P. Brooks, THE MYTHICAL MAN-MONTH 183 (1995). And that complexity may lead to unexpected results—which is, in fact, all we really mean by a 'bug'.

> computer files were last accessed, where evidence was introduced showing that computer was functioning properly and its clock was accurate]…..[52]

But the results of internal processing may use human-entered data as inputs, and that data of course can be challenged on a variety of grounds, including hearsay. Sometimes, data appears to come directly to the computer without human intervention, such as by way of sensors and digital imaging, in which case these may be thought of as part of the internal processing of the computer. There may be some difference between the standards used for the admission of the results of internal processing, as opposed to those used in connection with sensor input. In the former situation, as we will see below, the proponent must present some foundation (but not much) on the accuracy of the system, but the inputs of real time sensor information, such as automated photographs, "are presumed to be accurate,"[53] a test reminiscent of that applied to the 'print function' of a computer.

Turning then directly to classic internal processing, the basic rule for admissibility is succinctly captured in this unpublished opinion:

> The test for admissibility of machine created information is whether the computer was operating properly at the time of the printout. …. The admissibility of computer records also does not require establishing the accuracy, maintenance, reliability or the acceptability of the computer's hardware or software. (*People v. Martinez* (2000) 22 Cal.4th 106, 132.) Our Supreme Court has noted that mistakes can occur with computer generated information. However, such mistakes should not affect admissibility but be developed on cross-examination. (*Ibid.*).[54]

This "machine created information" or computer generated evidence (CGE) includes for example metadata, such as timestamps and document author identification information,[55] which, we say, is created automatically by the machine. The truth is that even for metadata, some human input is implicated such as by setting the time either on the machine itself or some other machine to which it refers, creating the author's initials, and so on. Nevertheless, the conceit here is that the data is as a general matter automatically created, and so qualifies as CGE. Thus no hearsay objection applies.

And while CGE does have to be validated with some foundational testimony, the bar is not high:

---

[52] 1 B. Witkin, CALIFORNIA EVIDENCE, Hearsay § 231 (5th ed. 2012).

[53] This is so "especially [for] government-maintained computers, [which] are presumed to be accurate. Thus, a witness with the general knowledge of an automated system may testify to his or her use of the system and that he or she has downloaded the computer information to produce the recording. No elaborate showing of the accuracy of the recorded data is required. Courts in California have not required "testimony regarding the ' "acceptability, accuracy, maintenance, and reliability of ... computer hardware and software" ' in similar situations. (*People v. Martinez* (2000) 22 Cal.4th 106, 132 …, quoting *People v. Lugashi* (1988) 205 Cal.App.3d 632, 642…[CLETS (California Law Enforcement Telecommunications System) printout]; accord, *People v. Nazary* (2010) 191 Cal.App.4th 727, 755 …)" (*Goldsmith, supra*, 59 Cal.4th at p. 272…[automated traffic enforcement system].)" *People v. Dawkins*, 230 Cal. App. 4th 991, 1003 (2014) (citations abridged).

[54] *People v. Johnson*, No. F069414, 2016 WL 4482963, at *3 (Cal. Ct. App. Aug. 25, 2016) (unpublished). See also, e.g., *People v. Hawkins*, 98 Cal. App. 4th 1428, 1449–50 (2002) ("the true test for admissibility of a printout reflecting a computer's internal operations is not whether the printout was made in the regular course of business, but whether the computer was operating properly at the time of the printout"). This foundational showing (that the computer was 'operating properly') doesn't require much. *People v. Martinez*, 22 Cal. 4th 106, 132 (2000) ("our courts have refused to require, as a prerequisite to admission of computer records, testimony on the 'acceptability, accuracy, maintenance, and reliability of ... computer hardware and software.'") (quoting *People v. Lugashi*, 205 Cal.App.3d 632, 642 (1988)). Mistakes can be exposed by cross examination. Id. Accord, *People v. Dawkins*, 230 Cal. App. 4th 991, 1003 (2014) (especially with government maintained computers, no elaborate foundational showing required); *People v. Peyton*, 229 C.A.4th 1063 (2014).

[55] *United States v. Hamilton*, 413 F. 3d 1138 (10th Cir. 2005).

> First, the witness through whom the computer records are introduced is qualified if that witness generally understands the system's operation and possesses sufficient skill and knowledge to properly use the system and explain the resultant data, even if the witness is unable to perform every task from initial design and programming to final printout. [Citations] Second, testimony on the acceptability, accuracy, maintenance, and reliability of computer hardware and software need not be introduced, particularly where the data consists of retrieval of automatic inputs rather than computations based on manual entries. [Citations].[56]

The requirement generally to explain the system's operation seems to satisfy, or be the functional equivalent of offering, "foundational evidence that the computer was operating properly."[57] What does it mean to show that the computer was operating "properly"? It seems to be no more than showing that it was operating as it usually does, explained by someone with some experience in using the system.[58] That's enough to meet the "minimum requirement for admissibility"[59] which, after all, still leaves the evidence subject to cross examination and argument that the fact finder should disregard it.[60]

3. *Simulations*

a. Foundation

The low threshold for CGE may simply be a practical response to an otherwise impossible problem. Computers and their data are of course ubiquitous, but no one really knows how they work in detail.[61] No

---

[56] Simons, above note 43 at § 2:63. See above, note 54. But for a lengthy list of issues concerning this functionality which conceivably might be up for discussion, see Gregory P. Joseph, "A Simplified Approach to Computer-Generated Evidence and Animations," 43 N.Y.L. SCH. L. REV. 875, 882 *et seq.* (2000). This list of issues may be significant as the focus of the (1) opposing side's attack on functionality, hoping to destroy the presumption of reliability (compare above, note 48), or (2) the proponent's efforts thereafter to shoulder the real burden of demonstrating functionality. If the evidence is admitted, the issues could also be used to argue to the fact finder that the evidence is or is not persuasive.

[57] *People v. Hawkins*, 98 Cal. App. 4th 1428, 1450 (2002). Cf., *People v. Goldsmith*, 59 Cal. 4th 258, 271 (2014) (among other factors court considered showing that "the evidence was properly received in the normal course and manner of Inglewood's operation of its ATES program").

[58] E.g., *People v. Johnson*, No. F069414, 2016 WL 4482963, at *3 (Cal. Ct. App. Aug. 25, 2016) (unpublished) (scanner evidence admitted just based on "testimony of how staff used the scanner. He said the scanner's information was from the night of the shooting because its data only lasted "a day or two" and he correlated the scanner with the video surveillance system"); *People v. Johnson*, No. B224491, 2011 WL 4436451, at *2 (Cal. Ct. App. Sept. 26, 2011) (unpublished) ("Stoltz testified that the software company who owned the loan software designed a special program to extract "miscellaneous" type transactions posted during a specific time period. The result of running the special program was a list of "miscellaneous" type transactions for the period in question. This is sufficient to create an inference that the computer program was working properly"). As observed in note 49, new federal rules of evidence will ease the admissibility of computer evidence. New FRE 902 (13) *seems* to apply to CGE (while 902 (14) applies to computer stored data), but the Committee Notes make it clear that only authenticity is established through the certification procedures of the rule, not reliability as such. "Similarly, a certification authenticating a computer output, such as a spreadsheet, does not preclude an objection that the information produced is unreliable—the authentication establishes only that the output came from the computer." <http://www.uscourts.gov/sites/default/files/supreme-court-package_0.pdf>.

[59] *People v. Lugashi*, 205 Cal. App. 3d 632, 640 (1988).

[60] See also, e.g., *People v. Nazary* (2010) 191 Cal.App.4th 727, 753–755, overruled on others grounds in *People v. Vidana*, 1 Cal.5th 632, 648 (2016) (test of admissibility of machine-generated receipts from automated gas station island pumps is whether "machine was operating properly at the time of the reading").

[61] Mark Meysenburg, INTRODUCTION TO PROGRAMMING USING PROCESSING 256 (2d ed. 2015) (regarding the typical fly-by-wire automated aircraft controls systems, "No single person on the face of the earth truly understands everything there is to know about the software that keeps the airliner flying."). See generally, Samuel Arberman,

person for example can report on the detailed instructions used by the most ordinary operating system, not to speak of the myriad interactions between operating systems and applications found in every business and most homes in this country. But the test is reasonable, because it meets what I suggest is a fundamental predicate, which is that notions of reliability in the legal world mirror those we use in the 'real' or ordinary world. We use the same reasoning to declare that business records are exempt from the hearsay rule: if the hearsay is good enough for a business to rely on, it is good enough for a jury. So too here: if an entity relies on the validity of CGE in its day-to-day work, a jury is justified in making the same assumption of validity. Expecting much more would exclude CGE from our courts.

But this reasoning does not quite extend to justify the admission of computer simulations, which are purpose made for a trial—these are bespoke CGE. The techniques used are not employed in the run-of the-mill business.[62]

Simulations and animation are not the same.[63] Animations are a sort of supporting evidence which serves only to illustrate other testimony, much as a drawing of car accident scene by an eyewitness serves to illustrate and explain the witness' testimony.[64] As demonstrative evidence, it may or may not be admissible,[65] but in any event it entirely depends on the primary testimony, and it is the human witness who is cross examined.[66] No one really cares how the animation was made, that is, how the (for example) drawing program works, or how it calculates distances or other feature, just as no one cares how a camera works when a witness testifies that picture is a fair representation of the scene she saw.

Simulations by contrast are introduced as primary or "substantive" evidence: they depend on accurate inputs, but their validity also, critically, depends on valid algorithms. The validity of the algorithm is fair game for challenge as it is not when an animation is presented:

> Courts have compared computer animations to classic forms of demonstrative evidence such as charts or diagrams that illustrate expert testimony. [Citations] A computer animation is admissible if " 'it is a fair and accurate representation of the evidence to which it relates....' " [citations] …. A computer simulation, by contrast, is admissible only after a preliminary showing that any "new scientific technique" used to develop the simulation has gained "general acceptance ... in the relevant scientific community.[67]

OVERCOMPLICATED: TECHNOLOGY AT THE LIMITS OF COMPREHENSION 3 (2016) (discussing in particular computer enabled systems, "technology complexity has eclipsed our ability to comprehend it.").

[62] This is not always true. Some businesses do indeed rely on simulations for their quotidian work. *Lapsley v. Xtek, Inc.*, 689 F.3d 802, 815 (7th Cir. 2012) ("simulation is one of the most common of scientific and engineering tools. Around the world, computers simulate nuclear explosions, quantum mechanical interactions, atmospheric weather patterns, and innumerable other systems that are difficult or impossible to observe directly. A mathematical or computer model is a perfectly acceptable form of test."). I speculate that these simulations might be admissible under a lesser lever of scrutiny, or that reliability might simply be easier to establish.

[63] *People v. Duenas,* 55 Cal. 4th 1 (2012). For a detailed discussion, see e.g., Curtis Karnow, LITIGATION IN PRACTICE at 39 *et seq.* (2017)

[64] E.g., *People v. Hood*, 53 Cal. App. 4th 965, 969 (1997).

[65] That is, judges may not let the animation go to the jury room during deliberations, and it may not become part of the record sent to the court of appeal. But of course the jury sees it, so in that less technical sense the animation is admitted.

[66] Betsy S. Fiedler, "Are Your Eyes Deceiving You?: The Evidentiary Crisis Regarding the Admissibility of Computer Generated Evidence," 48 N.Y.L. SCH. L. REV. 295, 299 (2004) ("the testifying witness must state that the CGE portrays the disputed subject matter fairly and accurately").

[67] *People v. Duenas,* 55 Cal. 4th 1, 20–21 (2012), quoting *People v. Kelly*, 17 Cal.3d 24, 30 (1976).

For this custom CGE, much more than the minimal threshold noted above must be presented. If challenged, the proponent must satisfy the much stronger test of showing the methods used by the software are justified by science, for example, by a showing "'that the facts and data upon which the simulation is based "are of a type reasonably relied upon by experts in the particular field,' that the simulation is 'the product of reliable principles and methods,' and that the supporting expert witness 'applied principles and methods reliably' when creating or using the simulation."[68] As one commentator notes, "in the context of simulations, the computer itself is the expert."[69]

Simulations may be made of, e.g., airplane crashes. The input consists of information such as records of radar returns, facts concerning the crash site such as distances between pieces of the aircraft, so-called "black box" data as speed over time, whether flaps were deployed, and so on; these are fed to a program which might reproduce a view of the accident from the perspective of the pilots, or provide a basis to say that such and such a warning must have sounded and been ignored, or that the aircraft was at a certain angle of attack. A simulation of ground water contamination might take inputs of data such as (i) measurements of a toxin over time in a certain area, (ii) the movement rate of groundwater over that period, and then in effect opine that the toxin must have been at a certain concentration at a point upstream at a specified earlier time.

In these situations the validity and hence admissibility of the simulation depends on the validity of the programming. Thus the California Supreme Court has held that simulations are admissible if they are scientifically reliable, and "only after a preliminary showing that 'any new scientific technique' used to develop the simulation has gained general acceptance ... in the relevant scientific community."[70] More established techniques must still be explained because, as expert systems, they are subject to the usual strictures. Those include, as suggested, basic scientific reliability,[71] and non-speculative connections between the conclusions (or output) of the opinion and the input.[72] But it is difficult to know what counts as sufficient demonstration.

b. Interlude: Explaining Software

Judges and juries expect that at some level the operations of simulations can be explained, that is, that the "heuristic basis" can be demonstrated.[73] For example, an expert might use commonly accepted formulae for the relationship between pressure of a liquid and the aperture of the container through which it is released, based on the classic Bernoulli equation, to compute the speed of the liquid or the pressure it would exert on its target.[74]

---

[68] Victoria Webster, Fred E. (Trey) Bourn III, "The Use of Computer-Generated Animations and Simulations at Trial," 83 DEF. COUNS. J. 439 (2016) (notes omitted) (includes multi-circuit survey).
[69] Op. cit.
[70] *People v. Duenas,* 55 Cal. 4th 1 (2012).
[71] *People v. Jackson*, 1 Cal. 5th 269, 320 (2016) (expert "procedures and experiments must comply with the laws of physics, chemistry, and biology"). See e.g., *Liquid Dynamics Corp. v. Vaughan Co.*, 449 F.3d 1209, 1221 (Fed. Cir. 2006) (simulations subject to analysis under classic *Daubert* criteria and deemed in this case to be reliable); *Novartis Corp. v. Ben Venue Labs., Inc.,* 271 F.3d 1043, 1054 (Fed. Cir. 2001) (valid "simulation …. requires both a solid theoretical foundation and realistic input parameters to yield meaningful results. Without knowing these foundations, a court cannot evaluate whether the simulation is probative"); *Lyondell Chem. Co. v. Occidental Chem. Corp*., 608 F.3d 284, 294 (5th Cir. 2010) ("we can gauge reliability by examining input values and requiring transparency from testifying experts").
[72] *Sargon Enterprises, Inc. v. Univ. of S. Cal*., 55 Cal. 4th 747, 771–72 (2012).
[73] William R. Swartout, "Explaining and Justifying Expert Consulting Programs," COMPUTER-ASSISTED MEDICAL DECISION MAKING at 254 (1985).
[74] *Lapsley v. Xtek, Inc.,* 689 F.3d 802, 815 (7th Cir. 2012).

> Computer-generated simulations are based on mathematical models, and particular attention must be paid to the reliability and trustworthiness of the model. A model is a set of operating assumptions — a mathematical representation of a defined set of facts, or system. To be accurate, it must produce results that are identical or very similar to those produced by the physical facts (or system) being modeled. In order to do that, the model must contain all relevant elements — and reflect all relevant interactions — that occur in the real world.[75]

But it is not clear what is involved in any foundational requirement to explain the operation of software, including computer simulations.

There are some issues of definition, such as whether the 'program' includes operating systems, interfaces, and commonly available libraries.[76] Those issues can to some extent be defined away. For example, judges probably don't want to hear anything about what we might call the general housekeeping functions such as the run time environment including operating systems, standard interfaces, and device drivers; and we probably don't care which language the program was written in (C++, FORTRAN etc.). These are aspects of base technology and ordinarily do not embody the decision making processes that are at issue as a foundation is laid for a simulation. For convenience, I will call the decision making mechanism of interest to the court the 'inference engine.' This is the part of the program that includes the manipulation of data and creates the conclusions.[77] It is the inference engine that embodies the central theories of the simulation, as opposed to more general theories of computation.

Setting issues of definition aside, more problematic concerns stem from the fact that software (including the inference engine) can be described at many levels of abstraction, down to what some call 'bare metal,' i.e. the machine code that executes on the central processing unit (CPU).

This problem of description is apparently handled in an ad hoc manner, for there is little useful authority. We do see suggestions that the foundation would include testimony "as to the accuracy of the equations used in the" simulation software,[78] or testimony on "a solid theoretical foundation and realistic input parameters,"[79] or that the proponent would "unravel his code and deduce the assumptions, algorithms, equations, and parameters that must be embedded within it," perhaps by "translat[ing] the foreign language of his computer model into a comprehensible language …."[80] Some courts ask for a showing that "the input and underlying equations are sufficiently complete and accurate …and …the program is

---

[75] Gregory P. Joseph, "A Simplified Approach to Computer-Generated Evidence and Animations," 1 SEDONA CONF. J. 55 (2000).

[76] A library contains 'prewritten' code with functions that can be called on by the main, executing program.

[77] The inference engine's code may, however, be located in a variety of subprograms and libraries. The in-court proponent of the software may or may not be cognizant of the specific mechanisms of each piece of the inference engine, because the engine might depend on components such as dynamic link libraries (DLLs) written by others, and the proponent may be wrong about what those DLLs do. Andreas Björklund, et al., "DLL Spoofing in Windows" <https://www.it.uu.se/edu/course/homepage/sakdat/ht05/assignments/pm/programme/DLL_Spoofing_in_Windows.pdf>. Ordinary programs built with so-called objective oriented programming (OOP) tools in effect hide their basic functionality within the 'objects' (components) sometimes built by others, Frederick P. Brooks, THE MYTHICAL MAN-MONTH 272 (1995), and, practically speaking, no witness is likely to be able to explain the processing of all these components.

[78] Laurie L. Levenson, DEMONSTRATIVE EVIDENCE—COMPUTER-DEPENDENT SIMULATIONS, CALIFORNIA CRIMINAL PROCEDURE § 22:26 (Rutter: 2016); accord, Edward A. Rucker and Mark E. Overland, CALIFORNIA CRIMINAL PRACTICE: MOTIONS, JURY INSTRUCTIONS AND SENTENCING § 48:22 (4th ed., March 2017).

[79] 3 WHARTON'S CRIMINAL EVIDENCE, Computer generated exhibits § 16:22 (15th ed.).

[80] *Novartis Corp. v. Ben Venue Labs., Inc.,* 271 F.3d 1043, 1054 (Fed. Cir. 2001).

generally accepted by the appropriate community of scientists".[81] None of this tells us exactly what sort of explanation is enough to lay a foundation.

We certainly do not want to be led step by step through the bare metal code, i.e., machine language[82] or, just a bit less raw, assembly language[83] which is understood by some (but not many) humans. Nor, I suggest, does the judge (or later, the jury) want to be led through the next higher level of abstraction such as we find in source code, which most programmers use to write software.[84] At a high level of abstraction, we might have general flow-chart diagrams; but while these might summarize the components and

---

[81] *Commercial Union Ins. Co. v. Boston Edison Co.,* 412 Mass. 545, 549, 591 N.E.2d 165 (1992).

[82] In machine code, each instruction executes directly on the computer's CPU. Here's an example:

```
8020  78
8021  A9 80
8023  8D 15 03
8026  A9 2D
8028  8D 14 03
802B  58
802C  60
```

Here's another example:

```
00000000
00000001
00000010
00000100
00001000
00010000
00100000
01000000
```

[83] This is an example:

```
Start:      .org $8020
   SEI
   LDA      #$80
   STA      $0315
   LDA      #$2D
   STA      $0314
   CLI
   RTS
   INC $D020
   JMP $EA31
   802D  EE 20 D0
   8030  4C 31 EA
```

[84] Source code looks like this:

```
static void
print_cookies(CURL *curl)
{
  CURLcode res;
  struct curl_slist *cookies;
  struct curl_slist *nc;
  int i;
  printf("Cookies, curl knows:\n");
  res = curl_easy_getinfo(curl, CURLINFO_COOKIELIST, &cookies);
  if(res != CURLE_OK) {
    fprintf(stderr, "Curl curl_easy_getinfo failed: %s\n",
            curl_easy_strerror(res));
    exit(1);
  }
```

processes of a system, they will not reflect most of the logical work or assumptions of the program. Those are *too* abstract.

Somewhere in between these levels of abstraction we have perhaps statistical formulas programmed into the inference engine, e.g., [85]

$$z_1 = \frac{\sum_{l=1}^{M} y_k^l \left(\prod_{i=1}^{n} \mu\, A_i^l (x_i)\right)}{\sum_{l=1}^{M} y_k^l \left(\prod_{i=1}^{n} \mu\, A_i^l (x_i)\right)}$$

Together with the program itself, formulas such as this are good candidates for disclosure to the other party (that is, to the other side's expert) because they (i) embody the statistical rules of the inference engine and (ii) in effect state the nature of the input and output. But on their own, they are obviously of no help to the judge or the jury. While it possible to explain such formulas in plain English, other forms of representation are more useful.

As one commentator has suggested this can be done in three ways: propositional logic, fuzzy logic diagrams, and decision trees.[86] In propositional logic, we can state the value of variables such as A and B, which can be true or false, but also can be one of any specified range of values. In the example provided by the commentator, A could have one of these there values: {BUY,HOLD,SELL}. Logical operators are applied, such as AnD, Or, or But Not, which are familiar to any lawyer who has done on line legal research. The results are predicates, which are then arrange to set out the program's strategy. Continuing with the example, we might have these variables as inputs: Price (P), Simple Moving Average (SMA), and Exponential Moving Average (EMA). The strategy might then look like this:[87]

IF (SMA > P) ^ (EMA > P) THEN ***BUY*** ELSE
IF (SMA > P) ^ (EMA < P) THEN HOLD

We might also employ fuzzy logic. That provides a *range* over which a variable is true, or belongs to a certain set. The programs might say conclude that a share of a company is a 20% BUY and 30% HOLD and 50% SELL. Inputs too can be expressed over a range, expressing degrees of uncertainly of their truth, which corresponds at least a high level with the way in which layers in neural networks decide whether or not to pass on a finding to the next layer. Diagrams can then be generated showing that when the combined input from a series of sources exceeds a threshold, a decision is reached. For example, a medical diagnosis system might have degrees of certainty and uncertainty concerning inputs such as {has headaches- to some degree}, {has a rash - to some degree}, {is nauseous- to some degree}, {has difficulty breathing- to some degree}; and then express a result {has Golem's Fever- with a certain level of certainty}.

Finally, decision trees show the impact of factors on a series of decisions, the basic structure of which is familiar to most people.

---

[85] This is a formula used in a fuzzy logic inference engine expert system. <http://what-when-how.com/artificial-intelligence/supervised-learning-of-fuzzy-logic-systems-artificial-intelligence/>.
[86] Stuart Reid, "10 misconceptions about Neural Networks," *Turing Finance* (May 8, 2014) <http://www.turingfinance.com/misconceptions-about-neural-networks/#blackbox>. Reid address neural networks, but his point is useful more generally.
[87] ∧ means 'AnD'; > means 'greater than' and < means 'less than'.

In flight simulators, for example, a series of models or subsystems such as the aerodynamic, gear, weather, and engine models (among many others) are inputs to equations which calculate motion, and those then in turn output to visual, sound, motion, instrument displays, and other outputs.[88] Each component model includes a series of equasions. For example, a sophisticated engine model will produce figures for "engine thrust, fuel flows and engine pressures and rotation speeds… engine failure modes (e.g. surge, stall or total failure).. [accounting for, e.g.,] engine characteristics [which] change considerably at low speeds and at very low altitude…."[89] The number of equations in a flight simulator is far beyond what could possibly be addressed at trial. So here too the practical approach will distinguish and then ignore aspects of the program which are routine and presumably generally accepted, from those which are novel. As to the latter, an expert can present a graphical representation of the decisions nodes,[90] the values at each which cause the node to make a decision one way or the other (e.g., "if [engine temperature] > [5000 degrees], output ['explode']"), together with a theory behind the figure, such as research that shows engines explode at certain temperatures.

There are two conclusions here. Importantly, for traditional expert systems a human expert's competence in demonstrating, explaining, and justifying the theory behind a calculation is crucial. Both the judge determining admissibility and the jury determining weight look to the human expert to vouch for the simulation[91] and explain step-by-step the way in which the software works, to state its assumptions, the valid scientific theories on which it is based,[92] and the logic used to spit out its results.[93] While opinions may be based on the results of programs, the human witness takes the credit, or suffers the impeachment, for the opinion. If the expert can't explain the model—how the software works and why it uses the numbers or formulas it does—then the evidence isn't admissible.

Secondly, there is a limit to explanation. All evidence at trial assumes other facts are true. We don't ask the contractor to prove her measuring tapes are accurate, or the doctor to prove the blood pressure cuff was accurate. While we may require as foundation for eyewitness testimony evidence the person was at the scene, we don't need testimony on how the eye and brain work to record and recall the memory recited in court. As noted above,[94] only a minimal foundation is required for the routine operations of computers. It is a waste of time to reinvent the wheel, as it were: routine operations are a given. So too with most of the foundation for the admissibility of simulations. We usually forego *all* explanation at the levels of greatest precision, for example, the source code level. We quickly pass by even high level descriptions of most of the calculations and built-in assumptions. At most we provide (i) high level

[88] David Allerton, PRINCIPLES OF FLIGHT SIMULATION 17 (Wiley 2009).
[89] Id. at 18.
[90] E.g., David Madigan, et al., "Graphical Explanation in Belief Networks," 6 JOURNAL OF COMPUTATIONAL AND GRAPHICAL STATISTICS 160-181 (1997) <https://pdfs.semanticscholar.org/f3c9/19412ab55ff2f5a34c26e8a727d0467243c3.pdf>.
[91] Elaine M. Chaney, "Computer Simulations: How They Can Be Used at Trial and the Arguments for Admissibility," 19 IND. L. REV. 735, 743 (1986).
[92] *In re TMI Litig.*, 193 F.3d 613, 669 (3d Cir. 1999), *amended*, 199 F.3d 158 (3d Cir. 2000). As the court noted, wonderfully summarizing the distinction in tests applicable to accepted versus new scientific theories, the "use of standard techniques bolster the inference of reliability; nonstandard techniques need to be well-explained." *In re Zoloft (Sertraline Hydrochloride) Prod. Liab. Litig.*, No. 16-2247, 2017 WL 2385279, at *6 (3d Cir. June 2, 2017) (note omitted).
[93] David Boies et al., "Computer generated evidence—Admissibility of computer simulations," ABA, BUSINESS AND COMMERCIAL LITIGATION IN FEDERAL COURTS § 66:17: (4th ed. December 2016 Update). See generally, *Novartis Corp. v. Ben Venue Labs., Inc.,* 271 F.3d 1043, 1051 (Fed. Cir. 2001) (requiring demonstration of the "assumptions made by [the expert] in his computer model, and ask whether they are supported by evidence in the record. These include both the theoretical principles that informed the model's design as well as the means by which its input parameters were derived.").
[94] T.a.n. 58.

explanations of a few central formulas, (ii) the foundation (studies, etc.) which justify those formulas, and (iii) to some extent the logic which links those two things. This is as it should be, given the constraints of time, the expertise of most judges and juries, and the essential task of trial which is relentlessly to focus on core material issues. But we should not delude ourselves: the trustworthiness of much evidence, including computer simulations, depends on a practically infinite network of unarticulated assumptions. Nevertheless, we say the evidence is reliable.

C. Admitting Machine Opinions

Humans can't explain how neural nets make their decisions.[95] But they can still establish that their results are reliable, because humans can explain (1) how the nets are trained, (2) how they were successful in the past, and (3) how they are successful with new data. These are the features which make nets reliable in the real world. Reliability in the real world is a sign that nets will be reliable in the courtroom. The three factors I have listed here are the basis for their admissibility in court.

I provide next arguments in favor of the reliability of machine opinion. First, we generally recognize and trust expertise the bases for which cannot be fully articulated, that is, tacit expertise. Second, we trust medications, sometimes with our lives, even though no one knows how they work. Third, neural networks are statistical models, and judges commonly rely on statistical models. Fourth, reliability is importantly a function of the ability to test and cross examine; and neural networks can, practically, be cross examined.[96]

1. *Tacit expertise*

In Gladwell's *Blink*,[97] an art expert regards a Greek statute offered to the Getty museum for $10 million. The expert declares it a forgery. He can't quite say why, but he's right. Much expertise is tacit: it cannot be well articulated. This is of course true in sports (how does a professional hit a baseball travelling at 100 m.p.h.?[98]), music, teaching, decisions by administrative agencies,[99] perhaps even judging,[100] and many other domains[101] where the expertise cannot be described although it can be observed. As opposed to

---

95 "'We can build these models,' Dudley says ruefully, 'but we don't know how they work.'" Will Knight, "The Dark Secret at the Heart of AI," MIT TECHNOLOGY REVIEW (April 11, 2017) <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>. Interestingly, Knight quotes a researcher, "It might just be part of the nature of intelligence that only part of it is exposed to rational explanation. Some of it is just instinctual, or subconscious, or inscrutable."

96 Jennifer L. Mnookin, "Repeat Play Evidence: Jack Weinstein, 'Pedagogical Devices,' Technology, and Evidence," 64 DEPAUL L. REV. 571, 577–78 (2015) (suggesting courts permit the "opposing party to replace given assumptions with alternative ones" to enable cross examination. "To be sure, we do not normally imagine that machine-generated evidence requires cross-examination, but it may be time to begin thinking in those terms") (note omitted).

97 Malcolm Gladwell, BLINK (2005).

98 The batter has about 125 milliseconds to decide, far less time than it takes to blink. This makes the task impossible. But batters use unspecific information from the movements of the pitcher *before the pitch* to estimate a likely pitch. Alex Kuzoian, "Hitting a Major League fastball should be physically impossible," *Business Insider* <http://www.businessinsider.com/science-major-league-fastball-brain-reaction-time-2016-4>.

99 Jacob Gersen et al., "Thin Rationality Review," 114 MICHIGAN LAW REVIEW 1255 (2016).

100 Lee Epstein, William M. Landes & Richard A. Posner, THE BEHAVIOR OF FEDERAL JUDGES 5 (2013); Chad M. Oldfather, "Of Judges, Law, and the River: Tacit Knowledge and the Judicial Role," 2015 JOURNAL OF DISPUTE RESOLUTION 155, 156 (2015) ("much of what goes into the process of decision-making is inarticulable") <http://scholarship.law.missouri.edu/cgi/viewcontent.cgi?article=1717&context=jdr>.

101 Let us not forget the work of high end travel agents. R. Buckley, "Decision making by specialist luxury travel agents," 55 *Tourism Management* 133-138 (2016)

novices who use explicit step-by-step processes, experts tend to use more conceptual structures to solve problems; but it is difficult to use these structures to actually explain the work to others.[102]

The inarticulate bases for some expert opinion presents a challenge to the usual way in which tests for admissibility are thought of. For example, under California's *Sargon* test, judges should be presented with the express logic of the reasoning between (A) the opinion and (B) its foundation, including (1) the facts of the case, and (2) the general theory and techniques used as demonstrably founded on studies or other sources.[103] That is, the test puts a high premium on the articulation of the connection or "logic" between, (a) studies and other foundations that establish a general theory or technique, and (b) those theories (or techniques) and the facts of the case, on the one hand, and the ultimate opinion, on the other hand. The reasoning or "logic" should be express. This allows the judge evaluating admissibility to determine that each step in the process is reliable.[104]

But this can't be entirely right. Experts with "special knowledge, skill, experience, training, or education" can testify,[105] and not all that experience or skill can be articulated. "Skilled experts" may have only many years of experience as a banker or landowner testifying to value of property, or years of experience as a carpenter, plumber, tile layer, and so on, and be perfectly good expert witnesses.[106] For some of these experts, there is only so much they can say about the foundation of their opinions, although we may perhaps expect "even … a witness whose expertise is based purely on experience, say, a perfume tester able to distinguish among 140 odors at a sniff, whether his preparation is of a kind that others in the field would recognize as acceptable."[107]

<http://s3.amazonaws.com/academia.edu.documents/46870648/Decision_making_by_specialist_luxury_travel_agents.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1498433718&Signature=QC%2B%2BB3s3xuTlr%2FyzEpyplr3NAqM%3D&response-content-disposition=inline%3B%20filename%3DDecision_making_by_specialist_luxury_tra.pdf>.

[102] Pamela J. Hinds et al., "Why Organizations Don't 'Know What They Know': Cognitive and Motivational Factors Affecting the Transfer of Expertise," *Sharing Expertise* 5 (ed. Mark S. Ackerman et al.) (2003) <http://s3.amazonaws.com/academia.edu.documents/34753733/sharing_knowledge.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1498431104&Signature=W1MhnFiqh1tUX5EXQyKySgYwPVI%3D&response-content-disposition=inline%3B%20filename%3DSharing_Expertise_This_page_intentionall.pdf#page=23> (regarding experts' "conceptual, abstract representations is that they appear to be simplified representations of the task. As experts begin to automate aspects of the task, details of the task become less salient and experts begin to view the task in an oversimplified way. In an experiment, Langer and Imber (1979) found that experts' lists of task components contained significantly fewer and less specific steps than did the lists of those with less expertise. Developing abstract, simplified representations of the task allows experts to process information more rapidly, view the task holistically, and avoid getting bogged down in details. As such, abstract and simplified representations generally serve experts well. However, there are situations in which these representations can interfere with experts' ability to share their expertise, particularly with others who have significantly less expertise").

[103] M. Simons CALIFORNIA EVIDENCE MANUAL § 4:22 (2015). See also, *Jennings v. Palomar Pomerado Health Systems, Inc.,*114 CA4th 1108, 1117 (2003) ("when an expert's opinion is purely conclusory because unaccompanied by a reasoned explanation connecting the factual predicates to the ultimate conclusion, that opinion has no evidentiary value because an expert opinion is worth no more than the reasons upon which it rests.") (internal quotes omitted). See generally my "Expert Witness: *Sargon* and the Science of Reliable Experts" in Curtis Karnow, LITIGATION IN PRACTICE at 161 *et seq*. (2017).

[104] E.g., *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 745 (3d Cir. 1994) (requiring "conclusions supported by good grounds for each step in the analysis ….[such that] *any* step that renders the analysis unreliable under the *Daubert* factors" is exposed).

[105] Evid. Code § 720(a).

[106] See Michael H. Graham, 5 HANDBOOK OF FEDERAL EVIDENCE § 702:6 at n.7 (7th ed. 2016) (lengthy list of occupations which qualify by virtue of experience).

[107] *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 151 (1999) as noted by Graham, above note 88, at n.24.

The results in two relatively recent cases, one from the California Court of Appeal and one from the Ninth Circuit, may be explained at least in part by the notion of tacit expertise. To the surprise of some trial judges (including no doubt the highly respected jurists reversed in these two cases), the trial courts' meticulous examination of the articulated foundations of the experts' testimony, which led to their exclusion, was set aside by the appellate courts. The appellate panels found that the trial judges had in each case glossed over the basic reliability of the opinion, as demonstrated by the experts' credentials and long experience.

In the state case, *Cooper*,[108] the trial judge had examined each study relied on by the expert, and found many problems. The appellate court however said that the expert, a cancer specialist, had looked at all the studies together and in his experience found them as a whole to be an adequate foundation. Importantly, I suggest, the appellate court went out of its way to cite, at length, the doctor's credentials and experience.[109] There were other issues *Cooper* had with the trial judge's approach, but the tenor of the opinion is that the witness was unquestionably an expert in the field, and if he found a basis for his opinions then the trial judge was in no position to second guess him.

In *Wendell*,[110] the Ninth Circuit too chastised the trial judge for "look[ing] too narrowly at each individual consideration, without taking into account the broader picture of the experts' overall methodology. It improperly ignored the experts' experience, reliance on a variety of literature and studies, and review of Maxx's medical records and history, as well as the fundamental importance of differential diagnosis by experienced doctors treating troubled patients." Here too, the appellate court gave a good deal of space to the experts' credentials and remarkable experience in the relevant fields, noting that the doctors used for their court opinions the same techniques they use in their quotidian work.[111] The court made this point: "Nothing in *Daubert*, or its progeny, properly understood, suggests that the most experienced and credentialed doctors in a given field should be barred from testifying based on a differential diagnosis."[112]

In both these cases, the trial courts' crusade to analyze each part of the foundation, and each step in the logical progression from foundation to opinion—although seemingly called for by state and federal supreme court precedent—foundered on the rock of the witnesses' more general expertise, measured by their credentials such as their education, years of expertise, experience in treating, list of publications, and the like.

These cases, and the fact that skilled experts' testimony is admissible even though it may be impossible to fully articulate the foundation for it, suggests that admissibility of expert opinion is often a function of the *general* reliability of the source, sometimes not so much the express articulation of the individuated reasons and steps taken in reaching the opinion. That is, we recognizes tacit expertise. And neural networks have tacit expertise.[113]

---

108 *Cooper v. Takeda Pharm. Am., Inc*., 239 Cal. App. 4th 555, 562 (2015).
109 *Cooper*, 239 Cal. App. 4th at 563.
110 *Wendell v. GlaxoSmithKline LLC*, __F.3d__,__, 2017 WL 2381122 (No. 14-16321, 9th Cir. June 2, 2017) at *4.
111 Id. at *5.
112 Id. at *6.
113 See generally, Jason Millar, et al., "Delegation, relinquishment, and responsibility: The prospect of expert robots," ROBOT LAW 102, 109 *et seq*. (R. Calo et al., eds. 2016). Millar and his co-author suggest that (i) much 'expert' knowledge is tacit, (ii) machine intelligence can or will manifest such knowledge, and (iii) we should defer to the conclusions of machine intelligence when it demonstrably does a better job than humans in given domains.

2. *The Drug Analogy*

Even when prescription drugs are approved, not all the side effects, or the factors on which they depend, may be known. So too with their benefits: Not all factors are known affecting the efficacy of a drug. Drugs are evaluated after they are first approved, and warnings may change over time. As the Food and Drug Administration (FDA) notes, "In the end, no matter how much data are available, we often have to make a judgment call, weighing the known benefits against known risks and the potential—and possibly unknown—risks."[114] But more to the point for present purposes, all that may be known about the approved drug is that it has certain benefits and certain other effects, without knowing the details of why that is so. That is, a drug may be approved for use, both over the counter and by prescription only, despite the fact that its mechanism is not known. Examples include "acetaminophen for pain relief, penicillin for infections, and lithium for bipolar disorder, [which] continue to be scientific mysteries today."[115] A 2011 study of 75 drugs found that only 17 were derived from a "detailed understanding of how the disease worked."[116] The point of course is that after sufficient trials, with a sufficiently representative population, the benefits and detriments of drugs may be sufficiently evident to allow them to be used, even if we don't know why they work.

In related contexts, we are comfortable with the fact that while we may not know the mechanism of injury, we know enough—generally through statistical studies—to find that a putative cause (such as a drug) indeed has a certain effect (such as a birth defect).[117] While there is a difference between accepting an expert opinion on causation not knowing the mechanism where (1) we have a demonstrated statistical basis, and (2) in the case of a neural network where we do not have a demonstration of the specific statistical basis, the truth is that neural networks are statistical analyses, and their reliability can be demonstrated through validation. I take this up next.

3. *The Statistical Framework & Explaining The Logic*

As the drug analogy demonstrates, we use statistics to make very serious decisions. Statistics are in widespread use in court, and indeed judges may take judicial notice of certain statistical facts.[118] Statistics are used to support and combat class certification decisions,[119] evaluate DNA evidence,[120] show and disprove racial disparity[121] and age discrimination,[122] prove Fourth Amendment violations,[123] are used in

---

[114] https://www.fda.gov/aboutfda/transparency/basics/ucm269834.htm.

[115] https://www.washingtonpost.com/news/wonk/wp/2015/07/23/one-big-myth-about-medicine-we-know-how-drugs-work/.

[116] David C. Swinney & Jason Anthony, "How were new medicines discovered?" *Nature Reviews Drug Discovery* 10, 507-519 (July 2011) <http://www.nature.com/nrd/journal/v10/n7/abs/nrd3480.html>; see also, Tanya Lewis, "Mystery Mechanisms," *The Scientist* (July 29, 2016) ("Scientists still don't know exactly how some commonly used drugs work"); "How Does Acetaminophen Work? Researchers Still Aren't Sure," 92 *Science* 31-32 (July 21, 2014).

[117] *Daubert v. Merrell Dow Pharm., Inc*., 43 F.3d 1311, 1314 (9th Cir. 1995), cited in *Wendell v. GlaxoSmithKline LLC*, __F.3d__, __(No. 14-16321, 2017 WL 2381122, at *6) (9th Cir. June 2, 2017)). In California courts, medical causation depends on expert testimony that there is a "reasonably probable causal connection" between the injury and alleged cause, *Jones v. Ortho Pharm. Corp.*, 163 Cal. App. 3d 396, 403 (1985), i.e., greater than 50% odds. See e.g., *Uriell v. Regents of Univ. of California*, 234 Cal. App. 4th 735, 746 (2015) (discussion of the state's 'more probable than not' test); *Cooper v. Takeda Pharm. Am., Inc.*, 239 Cal. App. 4th 555, 594 (2015) (same).

[118] *Envtl. Law Found. v. Beech-Nut Nutrition Corp.*, 235 Cal. App. 4th 307, 325 n.7 (2015).

[119] *Duran v. U.S. Bank National Association*, 59 Cal.4th 1 (2014); *Mies v. Sephora U.S.A., Inc.*, 234 Cal. App. 4th 967 (2015).

[120] E.g., *People v. Venegas*, 18 Cal. 4th 47 (1998).

[121] *Alston v. City of Madison*, 853 F.3d 901, 908 (7th Cir. 2017). *Paige v. California*, 291 F.3d 1141 (9th Cir. 2002).

[122] *Karlo v. Pittsburgh Glass Works, LLC*, 849 F.3d 61 (3d Cir. 2017).

[123] *United States v. Soto–Zuniga*, 837 F.3d 992, 1002 (9th Cir. 2016).

labor litigation,[124] to attack the practices at the Patent Office,[125] fix damages allocation in environmental clean-up actions,[126] and many, many other situations. Regression analysis is often used to estimate the impact of illegal acts.[127] Judges rely on statistical tools such as the STATIC-99 to evaluate the risk of recidivism of registerable sex offenders,[128] and many courts use the results of surveys and their statistical conclusions in sentencing[129] and to set bail, deciding issues such as the risk of recidivism and likelihood that defendants will appear for their next hearings.[130]

Neural nets are in effect statistical models.[131] A valid (or "statistically significant") statistical result shows a certain degree of correlation; it does not prove causation. But with traditional statistical studies, at some point either the strength of a study, or, better, many studies,[132] is enough to rely on, and to

---

[124] *Nat'l Labor Relations Bd. v. Lily Transportation Corp.,* 853 F.3d 31 (1st Cir. 2017).

[125] *Ethicon Endo–Surgery, Inc. v. Covidien LP*, 826 F.3d 1366, 1368 (Fed. Cir. 2016) (Newman, J,. dissenting from denial of rehearing en banc).

[126] *Lyondell Chem. Co. v. Occidental Chem. Corp.*, 608 F.3d 284, 292 (5th Cir. 2010).

[127] *E.g., In re Se. Milk Antitrust Litig.,* 739 F.3d 262, 285 (6th Cir.) *cert. denied sub nom. Dean Foods Co. v. Food Lion, LLC*, 135 S. Ct. 676 (2014); *Werdebaugh v. Blue Diamond Growers*, 2014 WL 2191901 (N.D. Cal. May 23, 2014) (proving damages under UCL, FAL, and CLRA); *Kleen Products LLC v. Int'l Paper*, 306 F.R.D. 585, 602 (N.D. Ill. 2015).

[128] <http://www.static99.org/>; see Penal Code § 290.03-08.

[129] Adam Liptak, "Sent to Prison by a Software Program's Secret Algorithms," *The New York Times* (May 1, 20-17), <https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html?hp&action=click&pgtype=Homepage&clickSource=story-heading&module=first-column-region&region=top-news&WT.nav=top-news> (describing Chief Justice Robert's apparent reference to risk assessment software used in sentencing, "'Can you foresee a day,' asked Shirley Ann Jackson, president of the college in upstate New York, 'when smart machines, driven with artificial intelligences, will assist with courtroom fact-finding or, more controversially even, judicial decision-making?' [¶] The chief justice's answer was more surprising than the question. 'It's a day that's here,' he said, 'and it's putting a significant strain on how the judiciary goes about doing things.'").

[130] See e.g., Peter J. Henning, "Is Deterrence Relevant in Sentencing White-Collar Criminals?," 61 WAYNE L. REV. 27, 38 (2015); Abstract, Martin Frisher et al., "Predictive Factors for Illicit Drug Use Among Young People: A Literature Review" (2007), <https://www.ncjrs.gov/App/AbstractDB/AbstractDBDetails.aspx?id=240320>; A. Kellermann, et al., "Gun Ownership as a Risk Factor For Homicide In The Home," 329 THE NEW ENGLAND JOURNAL OF MEDICINE 1084 (October 1993), <http://www.nejm.org/doi/pdf/10.1056/NEJM199310073291506>; T. Tillman, "Risk Assessment In Montana: Risk Factors Predictive Of Juvenile Offender Recidivism," Theses, Dissertations, Professional Papers. Paper 4495 (2015), <http://scholarworks.umt.edu/cgi/viewcontent.cgi?article=5421&context=etd>; Jason Tashea, "Bond Ratings Kentucky Is Testing A New Assessment Tool to Determine Whether to Keep Defendants Behind Bars, ABA J., 15, 17 (April 2015), < http://www.arnoldfoundation.org/new-data-pretrial-risk-assessment-tool-works-reduce-crime-increase-court-appearances/>; Curtis Karnow, "Setting Bail for Public Safety," 13 BERKELEY JOURNAL OF CRIMINAL LAW (2008); M. VanNostrand, et al., Pretrial Risk Assessment in the Federal Court. FEDERAL PROBATION 72 (2009), <http://www.pretrial.org/download/risk-assessment/Pretrial%20Risk%20Assessment%20in%20the%20Federal%20Court%20Final%20Report%20(2009).pdf>; K. Bechtel, et al., "Identifying the Predictors of Pretrial Failure: A Meta-Analysis," 75 FEDERAL PROBATION (2011) <http://www.cpoc.org/assets/Realignment/predictors.pdf>; Angèle Christin et al., "Courts and Predictive Algorithms," DATA & CIVIL RIGHTS: A NEW ERA OF POLICING AND JUSTICE (draft October 27, 2015) <http://www.law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf>.

[131] For a technical discussion, see R. Rojas, "Statistics and Neural Networks," NEURAL NETWORKS (Ch. 9) 229 *et seq.* (1996).

[132] Because individual studies may reflect cherry picking and other problems, studies which review the results of many studies (metastudies) are preferred. For discussion of a leading effort in this regard, see the Cochran systematic review <http://www.cochrane.org/cochrane-reviews>; <http://bmg.cochrane.org/addressing-reporting-biases> (Cochrane furthers transparency in research and publication, and use of metastudies); <http://community.cochrane.org/about-us/evidence-based-health-care/webliography/books/sysrev>;

conclude that given some predicate or sample, a more general conclusion is very likely to be true. Statistics use samples to tell us something about the world; they provide, based on partial data, an inference about general patterns. This how predictive coding works: Given a sample of documents on which the system has been trained, some of which are privileged (or relevant), which of the larger collection are privileged (or relevant)? AlphaGo works (in part) the same way: it extrapolates the present board pattern to an intermediate but limited set of possible patterns, given a series of possible moves; then it compares the intermediate set to all the patterns it has experienced.[133] Then, knowing which of that larger set led to victory, AlphaGo choses the best intermediate pattern and plays the move most likely to lead to it.[134]

There is of course a difference between the operations of neural nets and the traditional statistical expert, which is that the human expert 'shows his work'—we know how the calculation was done and we can inspect it for errors. We know if the formulas used are accepted in the relevant scientific community.

But we can also check the work of a neural net, even though we can't look under the hood. We *can* sort fallacious from reliable inferences. We do this by testing a neural network on new data. We validate the system. This sort of testing is how we figure out that a statistical correlation is invalid, i.e., that it is just a random product.[135] We can find correlations among almost any set of facts,[136] but this is cherry picking, and doesn't reflect a hypothesis which is then tested on new data. These untested correlations are thus—as we can often, but not always, tell from just looking at them—fallacious. They are random correlations selected after the fact—that's the cherry picking—because in isolation they appear to present a pattern.

But a fundamental element of training (whether human supervised or not) neural networks is a feedback loop which tests whether the learned correlations play out correctly on *new* data, just as networks used in business should have the predicted result constantly compared to real world events. AlphaGo was in this sense proved reliable: It beat the human champions. By the same token, networks can be tested against known results and a human witnesses can discuss, or challenge, the performance of a network against new data. The proponent of the machine opinion (say, for facial recognition or medical diagnosis), reports the performance of the network against new data and states that the program had correct results a certain percentage of the time. The party opposing admissibility, or (later) disputing the weight to be given to the opinion, could report results on his own set of new data. This process requires the software to be provided to all parties in order to allow for this sort of 'cross examination.' The selection of data by the party

---

<http://www.alltrials.net/> (importance of meta studies); Ben Goldacre, "Listen carefully, I shall say this only once," The Guardian (October 25, 2008), <http://www.badscience.net/2008/10/listen-carefully-i-shall-say-this-only-once/#more-823> (problems with issuing multiple reports of what is, in fact one study; contrasting results of the 'one' study with true metastudy results).

[133] This comprises *millions* of games, orders of magnitude more games than any human could play in a lifetime. "Full length games for Go players to enjoy," <https://deepmind.com/research/alphago/alphago-vs-alphago-self-play-games/>.

[134] D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," 529 NATURE 484 (28 January 2016) <http://www.nature.com/nature/journal/v529/n7587/full/nature16961.html>; Christof Koch, "How the Computer Beat the Go Master," SCIENTIFIC AMERICAN (March 19, 2016). <https://www.scientificamerican.com/article/how-the-computer-beat-the-go-master/>; D. Silver, "AlphaGo: Mastering the ancient game of Go with Machine Learning," Google Research Blog (Jan. 27, 2016) <https://research.googleblog.com/2016/01/alphago-mastering-ancient-game-of-go.html>

[135] A favorite example of fallacious induction is the Thanksgiving Day turkey who extrapolates from just under a year's worth of daily good feeding that Thursday November 23, 2017 will be a good day. It won't be.

[136] Such as (i) the number of people who drown by falling into a pool and films Nicolas Cage appeared in, or (ii) per capita cheese consumption and the number of people who died entangled in bedsheets; and so on and so forth. <http://www.tylervigen.com/spurious-correlations>. For more on bad or fallacious inferences, see my "Statistics & Probability: Bad Inferences and Uncommon Sense," in Curtis Karnow, LITIGATION IN PRACTICE 43 *et seq.* (2017).

opposing admissibility should be generated to detect flaws, such as training on inapposite data. I discuss this in the next section.

Furthermore, even though the technical calculation of the opinion is not available to any human, the proponent of the machine opinion should be able to provide to the layperson judge or jury an abstracted view of the logic flow that produces the opinion, similar to that available for a more traditional expert software with propositional logic, fuzzy logic diagrams, and decision trees, discussed above at § B(3)(b). The point here is not just that these three approaches can be used generally to illustrate machine decision making, but rather that there are tools to extract these illustrations *from the specific neural network* (i.e., its inputs and outputs) in issue.[137] While, again, these illustrations are not and cannot be descriptions of the actual mechanism of the hidden layers, nor justifications for them, they are probably as detailed as any descriptions provided to judges and juries in connection with more traditional inference engines.[138] That is, it often doesn't take much to provide as much detail as judges and juries really want, or need, in the evaluations of the proponent's foundation; because the real test for reliability is measured by the ability of the opponent to *challenge* the machine opinion. I turn to that next.

4. *Risks & Cross Examination*

The ability to cross examine is the classic test of reliability, and reliability is the cornerstone of admissibility.[139] I have already mentioned the *sine qua non* of cross examination, which is that the program must be made available to the opposing side in order to be tested against new data.

To a shocking degree, many ordinary 'expert' systems are in fact never tested against new data or real world results; their predications are never analyzed: they are never *validated*. They are used because they are convenient, they are cheaper than using humans, because they give the appearance of objectivity, or of infallibility, or to deflect responsibility from whoever would otherwise be the human agent, or to save time; or all of the above. Companies use software to hire and fire, but without figuring out down the road if the results were as predicted. Colleges use a variety of criteria to admit students, and use tests to measure competence in areas, and so on, but these may or may not have ever been validated—i.e. did the students with higher scores actually perform better in college? Did algorithms used to pick stocks actually do better than human decision-makers with the same information? As a matter of fact, what was the performance of loans when the lending decision was made by a program? Algorithms are used to suggest products on Amazon, movies on Netflix, plots for movies, patterns for room cleaning robots,[140] and for meeting people on dating sites online. Some of these *are* validated—especially, as in the case of companies such as Amazon and Netflix, because the accuracy of the algorithm spells millions of dollars.

---

[137] Stuart Reid, "10 misconceptions about Neural Networks," *Turing Finance* (May 8, 2014) <http://www.turingfinance.com/misconceptions-about-neural-networks/#blackbox>. Researchers continue to develop tools used to at least illustrate the details of training of specific networks. See e.g., Jason Yosinski, et al, "Understanding Neural Networks Through Deep Visualization," Deep Learning Workshop, 31st International Conference on Machine Learning (2015) ("We describe and release a software tool that provides a live, interactive visualization of every neuron in a trained convnet as it responds to a user-provided image or video") <https://arxiv.org/pdf/1506.06579.pdf>.

[138] See above, t.a.n. 93.

[139] This is true under both federal and California law. *Sargon Enterprises, Inc. v. Univ. of S. Cal.*, 55 Cal. 4th 747, 772 (2012).

[140] Kevin Slavin, "How algorithms shape our world," Ted Talk (July 2011) <https://www.ted.com/talks/kevin_slavin_how_algorithms_shape_our_world/transcript?language=en>

But many programs are never validated. The problem is sufficiently serious and pervasive that an entire book could be written about it.[141] Worse, at least from the point of view of those of us in the court system, many programs and tests used in criminal trials are of dubious validity because there are no accepted validation benchmarks, or no validation tests are used, or the level of precision announced to the jury is far, far in excess of the true value.[142]

So the first risk in using programs including neural networks is to note the lack of validation. And while I have pressed the notion that software which is relied on the world should usually be thought of as reliable in court, this is a critical caveat-- an exception to the rule that may swallow it. Unvalidated software is used all the time in the real world, but, as is true of any of the fallacies that infect our daily lives they have no place in court.

The critical importance of testing out a model, or any predictive system, on new data is exemplified by a nice example from Market Watch's Gary Smith.[143] He tells of his algorithm showing a remarkable 88% correlation between predicted and actual stock prices for 2015, including matching, almost perfectly, a drop in the third quarter:

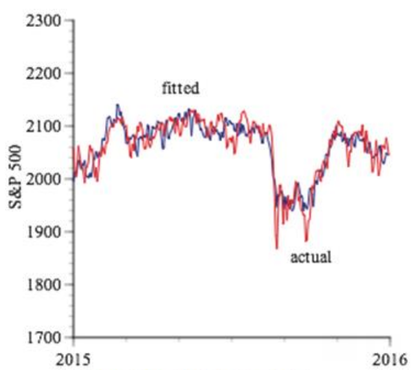


Figure 1 My Model of Stock Prices

But tested on new data—prices in 2016- it was a complete failure:

---

[141] Cathy McNeil, WEAPONS OF MATH DESTRUCTION (2016); see discussions of the issues, at e.g., <http://www.ibtimes.co.uk/dangers-big-data-how-society-being-controlled-by-mathematical-algorithms-1581174>. McNeil discussed the lack of feedback mechanisms (i.e. validation) in many areas, such as using algorithms to hire and fire teachers, id. at 7, 138, evaluating other potential employees, id. at 7, 111, issuing credit rating, id. at 146, and so on.

[142] President's Council of Advisors on Science and Technology (PCAST), "Report To The President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" (September 2016) <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf> (highlighting problems with tests re: certain DNA, bite-mark, firearms, hair comparison, fingerprint, and footwear (shoes). For serious problems with latent fingerprint testimony, see *U.S. v. Llera Plaza* 179 F.Supp.2d 492, 494 (E.D. Pa. 2002) *withdrawn from bound volume, opinion vacated and superseded on reconsideration*, 188 F.Supp.2d (E.D. Pa. 2002) (Pollak, J.). In 2015, the Washington Post reported that in almost all criminal trials for a period over 20 years in which a group of FBI finger print experts testified, including trials of 32 defendants sentenced to death, the FBI experts gave flawed testimony. Twenty-six experts overstated forensic matches in ways that favored prosecutors in over 95 percent of the 268 trials reviewed as of April 2015. <https://www.washingtonpost.com/local/crime/fbi-overstated-forensic-hair-matches-in-nearly-all-criminal-trials-for-decades/2015/04/18/39c8d8c6-e515-11e4-b510-962fcfabc310_story.html?utm_term=.e63ad6b8db16>.

[143] Gary Smith, "Opinion: This experiment shows the danger in black-box investment algorithms," (June 17, 2017) <http://www.marketwatch.com/story/this-experiment-shows-the-danger-in-black-box-investment-algorithms-2017-06-13\>

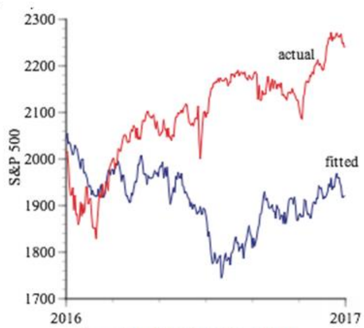


Figure 2 My Model with Fresh Data

And as with any validation, whether of a drug, a test, or some other screening device, we must ensure that the validation was conducted on the relevant population, on the relevant data. When testing drugs for childhood cancer, was the drug tested on 70 year olds with cancer because those subjects were more easily located? When testing an algorithm for recidivism and examining factors such as type of crime, income, or whether job history correlates with new crimes, or when looking at a program which predicts loan failures, was the validation population from the same type of locale – area of country, rural v. inner city—as the population on which the algorithm is to be used?

Underlying this issue is the problem of what *counts* as validation. The issue is simple with AlphaGo, because validation is obvious when it keeps beating every new opponent. (As it does.) So too with predictive coding of documents: we can examine the program's decisions on new data, and score accuracy: in these cases, the selection of a 'new' population of items used for validation is easy. It is also relatively easy to select items for medical imaging: we pick a series of x-rays or CT scans and score accuracy. But it is less certain what the new data (used for validation, which I will term 'validation data') might be for systems designed, for example, to do handwriting analysis, facial recognition, medical diagnosis, or recidivism.

For facial recognition systems, validation data might include photographs taken under a variety of lighting conditions, and of various angles of the face some of which will reveal few of the facial features. 'Successful' testing is likely to depend on which data we use. Similarly with handwriting analysis: validation data might include a wide variety of legible and illegible scrawls, initials, small and large groupings of letters; and 'success' is likely to depend on which of these are used. Medical diagnosis too depends on an unconstrained or arbitrary number of inputs, from a few to a very large number, such as body temperature, blood chemistry, as well as a range of vaguely reported conditions such as nausea, pain, skin tone, and extent of bruising. Recidivism may depend on different factors (a) in different areas of a country, and (b) such as different socio-economic groups. In all of these situations, and doubtless others, 'success' with one group of validation data may or may not be persuasive.

More technically with respect to neural networks, the validation data must meet certain criteria, such as that it not be the data used in the training.[144] Further 'cross-validation' against a much larger set of data can then correct errors in the initial validation.[145]

With programs that have extensive past experiences (i.e. they have worked on exceedingly large sets of data in the past) and are tested (including cross validated) on very large sets of validation data, these

[144] DEEP LEARNING at 118.
[145] Id. at 118-119.

concerns will tend to dissipate, just as the fact that AlphaGo has played millions of games obviates concerns that it may not be successful in the next game against a top professional player. The extent of the training and of the validation data in effect tells us that the next test—the one for which it provides the opinion in issue—is not 'unexpected,' not an outlier. It is no coincidence that neural networks have made their mark just as 'big data' erupted: it is commonplace to remark that we (or at least the companies that store it) have access to stunning amount of data,[146] as a function not only of efforts to digitize past records, but also the recordation of communications such as email, texts, searches, and social media which have taken the place of unrecorded oral communications (or, in the case of searches, unrecorded gestures such as picking up a dictionary or other reference sources) of the past. The accumulation of this data has not only made it imperative that we have software capable of digesting it, but is the very basis for the tools—neural networks—needed to do so. The stunningly large data set makes it reasonable to trust validation tests which use that data, all without knowing why it is that the validation is successful, that is, without either having a theory of correlation nor knowing the details of the underlying mechanism which explains the found correlation.[147]

There are two other dangers in the use of neural networks, which should be the subject of review by parties opposing the use of machine opinion. The first is closely related to the issue discussed above, that is, the make-up of the validation data. The issue is the use of proxies. Assume we use a network to opine on the relationship between a series of factors and fluctuations in fish stock.[148] A deeper review of the system might reflect the designer's decision not to measure the fish stock as such but some proxy for it, such as fish caught or consumed. Or a system designed to opine on earthquake damage might use simplified input of proxies of certain soil conditions.[149] A network might provide an opinion on the valuation of initial public offerings, but actually use a proxy such as valuation of certain stock one day or one week after the initial offering date. These may all be reasonable, but the underlying assumptions should be made manifest; sometimes, it may indicate a mismatch between the training data and the validation data on the one hand, and the proposed input for the specific opinion at issue, on the other hand.

Finally I refer to bias. The cartoon conceit is that algorithms are unbiased; the computer is neutral, free of prejudice. Without programming, the empty computer surely is. But most neural networks, even those which end up (as did AlphaGo) improving with self-training, begin their existence trained by humans; they literally model themselves on human choice and predilection. Some of these systems are in effect

---

[146] David Weinberger, TOO BIG TO KNOW (2011). Any measure of currently available data exceeds the mind's ability to understand, but for those of us who love words like *petabytes* and (even better) *zettabytes*, there is this: "The total amount of data in the world was 4.4 zettabytes in 2013. That is set to rise steeply to 44 zettabytes by 2020. To put that in perspective, one zettabyte is equivalent to 44 trillion gigabytes." Every day about 2.5 exabytes are produced, equivalent to 250,000 Libraries of Congress. Mikal Khoso, "How Much Data is Produced Every Day?," *Level* (Northeastern University, May 13, 2016) <http://www.northeastern.edu/levelblog/2016/05/13/how-much-data-produced-every-day/>. See the interesting predictions by Cisco, which as a manufacturer of internet servers and related equipment, presumably should know: "The Zettabyte Era: Trends and Analysis," (updated June 7, 2017) <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html>. Here's one my favorites: "It would take more than 5 million years to watch the amount of video that will cross global IP networks each month in 2021." Id.

[147] See the provocative Chris Anderson, "The End Of Theory: The Data Deluge Makes The Scientific Method Obsolete," *Wired* (June 23, 2008) <https://www.wired.com/2008/06/pb-theory/>.

[148] Compare e.g., D. G. Chen, "A neural network model for forecasting fish stock recruitment," 56 *Canadian Journal of Fisheries and Aquatic Sciences* 2385-2396 (1999), https://doi.org/10.1139/f99-178'; abstract at <http://www.nrcresearchpress.com/doi/abs/10.1139/f99-178>.

[149] C. Salameh, "Estimation Of Damage Level At Urban Scale From Simple Proxies Accounting For Soil And Building Dynamic Properties," *Proceedings of the 16th World Conference on Earthquake Engineering* 2049 (January 2017) <https://hal.archives-ouvertes.fr/hal-01461198/document>.

told that success is doing things the way humans would, and failure is diverging from those human choices. In this way, human biases become embedded in the very fabric of the system's decisions. The impact may be the most significant in what appear to be complex, subjective decisions such as hiring, evaluating written essays, and face recognition.[150] An interesting study of the way 30,000 images were used to train networks to recognize content found that human stereotypes on gender and race—i.e. prejudices—were routinely derived from the human tagged dataset.[151]

D. Conclusion

The results of well-trained neural networks are trusted in the world, and they can be trusted in court. Perfection is not guaranteed,[152] but it is not guaranteed with the testimony we routinely accept, such as eyewitness evidence[153] and confessions[154] which sometimes have peculiar reliability problems, and other routine testimony which maybe false, or mis-remembered.[155]

General admissibility rules are not meant to be onerous. The default is that all relevant evidence is admissible,[156] and if an opinion is reliable and relates to a contested fact, it is surely relevant. The foundations of expert testimony typically must be explained to the judge and jury, but this note demonstrates that in the case of opinions generated by neural networks, the fact that the specific basis for the opinion cannot be demonstrated or articulated should not block the admission of the opinion, because the opinion may yet be fundamentally reliable, and remains subject to meaningful cross examination. It may be that in the first case or two in which true machine opinion is offered, a so-called *Kelly* hearing may be warranted, because a court may find that the neural network is in this context an "unproven technique or procedure [used] … to provide some definitive truth which the expert need only accurately recognize and relay to the jury."[157] To avoid presenting the jury with a "misleading aura of certainty,"[158]

---

150 Joy Buolamwini, "How I'm fighting bias in algorithms," *TED Talk* (November 2016) <https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms/transcript?language=en>. See generally, Claire Miller, "When Algorithms Discriminate," *The New York Times* (July 9, 2015) <https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html>; Nanette Byrnes, "Why We Should Expect Algorithms to Be Biased," MIT TECHNOLOGY REVIEW (June 24, 2016) <https://www.technologyreview.com/s/601775/why-we-should-expect-algorithms-to-be-biased/>.

151 Emiel van Miltenburg, "Stereotyping and Bias in the Flickr30K Dataset," *Proceedings of the Workshop on Multimodal Corpora: Computer vision and language processing* (May 24, 2016) <https://arxiv.org/pdf/1506.06579.pdf>. See also, "Human Prejudices Sneak Into Artificial Intelligence," *Neuroscience News* (April 14, 2017) <http://neurosciencenews.com/artificial-intelligence-human-prejudice-6411/>, discussing Aylin Caliskan, et al., "Semantics derived automatically from language corpora contain human-like biases," *Science* (April 14, 2017).

152 DEEP LEARNING at 193.

153 Cross-racial witness identification may pose serious problems. See New Jersey's approach, *State v. Henderson*, 208 N.J. 208, 267, 27 A.3d 872, 907 (2011) *holding modified by State v. Chen*, 208 N.J. 307, 27 A.3d 930 (2011); *State v. Romero*, 191 N.J. 59, 68, 922 A.2d 693, 698 (2007). See *United States v. Langford*, 802 F.2d 1176, 1182 (9th Cir. 1986).

154 E.g., *People v. McCurdy*, 59 Cal. 4th 1063, 1109 (2014); *Campos v. Stone*, 201 F. Supp. 3d 1083, 1099 (N.D. Cal. 2016). See generally, Steven A. Drizin & Richard A. Leo, "The Problem of False Confessions in the Post-DNA World," 82 N.C. L. REV. 891 (2004); Welsh S. White, "False Confessions and the Constitution: Safeguards Against Untrustworthy Confessions," 32 HARV. C.R.-C.L. L. REV. 105, 119 (1997) (standard interrogation guidelines may induce false confessions).

155 Cf., *Trear v. Sills*, 69 Cal. App. 4th 1341, 1345 (1999); F. Lee Bailey et al., Fallibility of Memory, 2 CRIMINAL TRIAL TECHNIQUES § 58:11 (June 2017).

156 Evidence Code § 350.

157 *People v. Jackson*, 1 Cal. 5th 269, 316 (2016), quoting *People v. Stoll*, 49 Cal. 3d 1136, 1155–56 (1989).

the court may examine the technology. It may find that the basic technology is sound, widely used and accepted in the real world, and reliable; that correct scientific procedures (e.g., accepted statistical algorithms) were used to build and train the program, and that the network is helpful to the jury. With all parties well informed and able to validate functionality, machine opinions may provide insight no human can offer.

✠

[158] *People v. Kelly*, 17 Cal. 3d 24, 32 (1976), quoting *Huntingdon v. Crowley*, 64 Cal. 2d 647, 656 (1966). See generally, M. Simons, CALIFORNIA EVIDENCE MANUAL §§ 4:27 *et seq*. (2015).