Keynote Address on the Ethics of Artificial Intelligence
by Justice Gordon Goodman

UC Law Business Symposium — Corporate AI Toolbox
Academe Auditorium, UC San Francisco Law School
Friday, April 5, 2024

Book VII of Plato's Republic[1] opens with the strange image of individuals who live in a cave, who only see shadows on the walls of the cave, and who hear echoes that reflect what is happening outside the cave. To them, the truth would be "… literally nothing but the shadows of the images." For these cave dwellers, the process of turning to look at the light of truth outside the cave would be painful, and they would suffer. Because of this pain, the cave dwellers might think that the shadows they formerly saw were truer than the bright reality they now perceived.

Socrates argues that the change from viewing shadows to seeing the truth cannot happen in a moment, but that the former cave dwellers must be introduced to the good gradually. To do otherwise would lead to resistance from the cave dwellers who might think it better to not even think of ascending. He then notes that in the world of knowledge, "… the idea of the good appears last of all and is seen only with an effort…."

Background on Complexity Science and Artificial Intelligence

Let me start with some background on how we arrived at this critical moment, this tipping point, in the application and development of artificial intelligence, also called machine-aided learning, for use in the legal and corporate environment.

Looking first at my own experiences, I hope you will have a chance to review my most recent article titled "The Ethics of Artificial Intelligence" (about to be published in the Summer 2024 UC Law Business Journal) in which I describe my efforts during the late 1990's to develop a primitive version of artificial intelligence, based on techniques and approaches relating to complexity studies that were developed at the Santa Fe Institute in New Mexico during the 1980's and 1990's. This early AI application was intended for use in the energy industry.

---

[1] PLATO, THE REPUBLIC 233-239 (Benjamin Jowett trans., Teachers' and Students' Handy Edition, Books, Inc. 1943)

One point not mentioned in the article was that the success of the primitive AI program we developed twenty-five years ago was attributable in part to the fact that we could see emerging trends and complex patterns the other market participants, who were using traditional linear forms of analysis, could not see at that time. We were not dealing with a truly level playing field in which all market participants have access to similar AI-driven technology.

Another point not mentioned in my article is that small changes in the search parameters, the training periods utilized, and the datasets selected can all result in fairly large changes in the outcomes and recommendations of artificial intelligence programs. This is a word of caution for the future, and I will discuss this in more detail later in this address.

And for those of you who are unfamiliar with complexity science and its relationship to generative AI, here are some basics. In broad terms, there are two types of problems that we typically confront in our work. The first are the mechanistic but complicated problems that are solvable with linear or partially linear forms of analysis – including algebra, geometry, watchmaking, and bridge building. And then there are the organic, complex problems that require solutions derived from nature and evolutionary theory – including weather prediction, the movement of financial markets, politics, and ecology.

To address this second type of problem, and these are often the most interesting problems, the scientists of the Santa Fe Institute worked on solutions that mimicked the methods seen in nature (like evolution) to develop non-linear, complex approaches including most notably "genetic algorithms" and "neural networks" to solve problems that are otherwise unsolvable using traditional linear forms of analysis. This second group of problems often exhibit sudden changes – booms and busts – that are seemingly hard or impossible to predict.

Solutions from Complexity Science and Artificial Intelligence

The processes used within the first group of complicated but non-complex problems to optimize or improve results can sometimes bring complex situations to the very brink of failure. An example of this process involves the design of airplane wings. As you make wings more and more successful with regards to lift and speed, you are also approaching ever closer to a point of failure where there is no lift at all. This is sometimes referred to as a failure horizon.

Some of the solutions to complex problems that you see in our world today include "neural networks" (which is another technique developed by complexity science). These neural networks are used to mimic the behavior within our brains, and they are usually arranged into layers of virtual neurons, with information passing from the first input layer thru several intermediate layers (the hidden layers) to the final output layer. This is a basic tool in AI to identify patterns and make predictions.

The "internet," originally called the ARPANET and which some researchers consider to be a type of super neural network, was designed by DARPA (the Defense Advanced Research Projects Agency) to survive possible future attacks on major communication centers by moving

information not only down easily attacked central data corridors but instead in multiple directions from each server to all the other servers on the web in the same way that our brains move information diversely through a huge number of neurons and neural interfaces.

So, in looking to mimic human intelligence in designing artificial intelligence, complexity science was employed to solve the many complex problems human beings must confront. One technique in particular, the "genetic algorithm," is employed to find solutions that must evolve and change with changing inputs and changing circumstances over time. Inside a genetic algorithm program, generations of virtual agents look to solve complex problems and the best and most successful survive and combine to create the next generation of virtual agents. This can be conducted over hundreds or thousands of generations within these programs. Very often the solutions identified are novel and startling, but they are also often surprisingly effective.

One example of the use of genetic algorithms to solve a complex problem, in my case predicting the larger price movements of commodities seen in a global marketplace, is recounted in my new article. Though generative artificial intelligence uses many other techniques, at its heart there is a line of development from Santa Fe to the most recent AI programs being released in recent years.

The current process of artificial intelligence becoming suddenly popularized reminds me of the moment in the 1980's when the very useful but visually unattractive original software developed to drive personal computers, MS-DOS with its odd commands and computer language, was supplanted by the more attractive and user-friendly but functionally indistinguishable graphic user interfaces (GUI's) that arrived with the introduction of Lisa, MacIntosh, and Windows software. Similarly, today we are seeing the rapid success of attractive AI programs that rest on the underpinnings and logic derived from complexity studies conducted in Santa Fe thirty years ago.

My forward looking and optimistic early efforts to employ artificial intelligence were placed on hold by the burgeoning credit crisis that started with the Enron bankruptcy in late 2001 and reached its culmination with the Lehman bankruptcy in 2008. Beginning in 2001, I turned my focus to more defensive efforts that lead to several proposed changes in accounting and regulatory oversight through my work in Norwalk, CT, with other lawyers and accountants at the Financial Accounting Standards Board (FASB).

My efforts during this period are described in more detail in a series of articles I published over the last five years including "The Ethics of Finance," "The Ethics of Governance," and "The Ethics of Cryptocurrency" (the first published in the ABA Business Law Today, while the second and third were published by the UC Law Business Journal).

Human Lawyers and Virtual Associates

Let's turn now to the new relationship that will arise between lawyers and "virtual" associates. In the future, each of the young human lawyers in this room will employ thousands of "virtual" lawyers to assist in preparing contracts, drafting briefs, and (as I hope many of you will consider becoming judges later in your careers) writing opinions. The productivity improvements will be enormous and profound.

The partner/associate model will be the starting point for a new human/virtual relationship – human lawyers will have to guide and oversee the work of their "virtual" associates with the same care that partners currently give to their human associates. And just as bright associates can surprise and delight their partners one day but disappoint and distress their managing partners the following day, "virtual" associates will exhibit the same pleasing and disturbing behaviors, which will require constant vigilance.

How to Integrate AI technology in the Practice of Law?

Every parent knows the answer to this question. Because I worked for large corporations earlier in my career, I was forced to learn and adopt daily computer use while many of my peers delayed or deferred this important transition for years or decades. But my children did not have to learn or adopt anything, it was part of the fabric of their lives from day one, and they took to daily computer use as part of the air they breathed. It was natural and easy for them to make this dramatic societal change while it was often hard and painful for their parents.

So, just as my children patiently explained the intricacies of iPhone use and Facebook protocols to me and my wife, the law students sitting here in San Francisco will have to patiently explain the uses and best applications of AI technologies to their senior partners and employers for many years to come. In the same way that my work with intelligent businessmen and businesswomen twenty years ago to adopt a primitive form of artificial intelligence, as described in my new article for the UC Law Business Journal, was both frustrating and exciting, you will be challenged by the confusion and the fear that the new AI technologies will bring to the practice of law.

A key takeaway during any technological transition is to personalize the information and the work products that can be generated or derived from AI technologies. By this I don't mean we should simply place our names on AI generated documents or research, but rather that the final presentation, the final contracts, must always reflect your own final human oversight, review, and ultimately your ownership while acknowledging the input and support that you have derived from AI technologies.

An interesting development in "humanizing" AI technology was recently reported in a Washington Post article by Mr. Josh Tyrangiel about a version of ChatGPT that was modified and adapted for use by educators with input from Mr. Sal Khan and his team at the Khan Academy. Mr. Khan determined that though the version of ChatGPT he started using was good at computations, it was also easily bullied. When presented by a human user with a clearly inaccurate calculation, that earlier version of ChatGPT would often quietly and submissively

agree. To counter this type of "hallucination," the lesson plans, essays, and sample problems from Khan Academy were then downloaded into the ChatGPT database. Based on this feedback, they were able to train the next version of ChatGPT to become a more patient and knowledgeable virtual "teacher."

Like the example of training a virtual "teacher," the law students with me here at the Symposium will need to train their virtual "lawyers" and virtual "legal assistants" not to back down when confronted with inaccurate or misleading assertions even when made by experienced human lawyers.

Artificial Intelligence as a Logical Extension of the Legal Indexing System

Before the introduction of computers used by lawyers, beginning in the 1960's, the legal indexing/researching system was the largest and most complex manual dataset ever created. Young law students and lawyers were required to learn how to Shepardize and check citations against physical law books and reporters that were updated monthly, often with loose-leaf binders.

Starting in the 1970's, lawyers began to use computer based legal indexing/researching systems to cite check draft briefs and opinions instead of manual searches. Adoption of the computer aided Lexis and Westlaw systems proceeded throughout the latter part of the 20th century. Though manual cite checking continued for several decades, this transition was largely completed by the end of the 20th century and early years of the 21st century.

In the 2020's, lawyers will begin to use artificial intelligence software not only for cite-checking and proof reading (capabilities that already existed in pre-AI software programs used by many lawyers) but also for drafting and analysis at the front end of legal writing. Also, in the 2020's, lawyers will begin to use artificial intelligence not only for the selection of existing forms and model contracts (capabilities that have been available to some extent for over a hundred years) but also for preparing and drafting complex contracts for novel and previously undocumented transactions.

Like the earlier computer aided research transition of the 1970's to early 2000's, the transition to adoption of artificial intelligence in the legal profession will probably continue for several decades across the U.S. and around the world.

Licensing and Certification of Artificial Intelligence "Virtual" Lawyers

This brings us to the present. Lawyers and law firms may soon choose to employ AI agents as "virtual" lawyers to help draft and analyze legal issues and to help write legal contracts, draft briefs, and prepare opinions. But as with law students who have recently graduated from law school but who have not yet passed the bar exam, "virtual" lawyers used for these purposes

must prove their compliance with all bar association requirements and applicable legal professional Codes of Conduct and related ethical and moral principles.

As many of you know, a bill was introduced by Senator Scott Wiener (who represents San Francisco in the California Senate) on February 8, 2024, that would require companies training new AI models to test their tools, including presumably AI software prepared for lawyers, for any "unsafe" behaviors. With regards to the legal profession, we have already seen some indications of the "unsafe" behaviors that can arise when AI models draft briefs that cite fictional cases with real-looking references – known as AI hallucinations.

Several cases have been reported within the last year, including the Mata v. Avianca, Inc. case. In that case, the attorney who used ChatGPT to write a brief noted that he was under the erroneous impression that ChatGPT was a type of search engine. He claimed he was shocked to learn that ChatGPT was capable of fabricating real looking case citations. In another case, Park v. Kim, a lawyer was referred to the attorney grievance panel after she used ChatGPT to help with research in a medical malpractice case but unfortunately did not confirm the citations suggested by AI were valid.

Perhaps the most interesting recently reported use of ChatGPT was to support the proposed attorney's fees in a New York case filed under the Disabilities Education Act. The judge overseeing the case noted that ChatGPT did not identify the input data used in reaching its conclusion about a reasonable rate and as a result it was impossible to know whether the output suggested rate was real or relevant.

Another key takeaway is that even when using "virtual" lawyers who have passed their "virtual" bar exams, real human lawyers will still retain the ultimate responsibility for their filed briefs and opinions and for the contracts they prepare for clients.

Artificial Intelligence as the Basis for a New Approach to Legal Analysis

I will leave the discussion about what the future will bring technologically to some of the great experts you will be hearing from later today, but here are my thoughts about a new approach to legal analysis that may arise in the 21st century based on inputs and ideas from the new artificial intelligence software.

Justice Oliver Wendell Holmes, reacting to the upheavals during and after the Civil War, argued that courts must consider modern circumstances when interpreting the constitution and laws of the United States. He pointed back to the Common Law system of jurisprudence in which change was gradual but steady over long periods of time but also addressed the needs of the present – the "felt necessities" of our times. His study from 1886 titled the "Common Law" influenced Brandeis, Frankfurter, and several generations of Supreme Court justices. The Common Law approach to judicial analysis is part of the larger intellectual movement in moral philosophy called Pragmatism. This is the approach that is closest to my own judicial philosophy.

In reaction to the Common Law Approach, starting in 1971, Judge Bork and Justice Scalia argued for a return to Originalism -- a return to the original meaning of the text as written at the time of adoption. With the recent appointment of mainly originalist justices to the Supreme Court, Originalism or textualism has largely become the modern rule (with some variations for plain language approaches).

All previous transitions in mainstream legal analysis have looked to the past or the present for guidance. The Common Law looked to the present, and Originalism looked to the past. It's interesting that neither of these approaches significantly looked to the future. This was in part because of limitations on making future predictions. My question for this Symposium is whether the inherent predictive capability of Artificial Intelligence signals that the next major transition in legal analysis will include future outcomes?

I will leave the answer to this important question to the young lawyers and law students in this room.

Justice Gordon Goodman was elected to the First Texas Court of Appeals in 2018. He is a member of the Texas State Bar, Pennsylvania Bar, and Energy Bar Associations. Before serving on the bench, he started his career with the Whittenburg Law Firm in Amarillo, TX, where he focused on civil trials, appellate work, oil & gas law, banking law, and general practice. He subsequently served in senior positions for the Howell Corporation; E.I. DuPont de Nemours & Co.; Conoco, Inc.; Occidental Petroleum Corporation; and most recently with NRG Energy. He earned his BA degree Magna Cum Laude from the University of Pennsylvania and his JD degree from the University of Pennsylvania Law School both in Philadelphia. He received his high school degree from the Horace Mann School in the Bronx, NY.