AGENCY RFI RESPONSE

To NIST's Assignments Under Sections 4.1, 4.5 and 11 of the Executive OrderConcerning Artificial Intelligence (Sections 4.1, 4.5, and 11)

- **ID NIST-2023-0309**

MANAGING AND REGULATING THE TECHNOLOGY OF ARTIFICIAL INTELLIGENCE

Myron S. Karasik, MSEE, MBA, CPA, CMC, CTO & CFO emeritus (multiple)
McCracken Institute, Atlanta GA

There is much angst over the emergence of 'generative AI' with both utopian and dystopian visions abounding. Having been involved with this technology for over a half century, learning it from the founder of the first generation of this technology used in the LANDSAT land management program and also used for chromosomal anomaly detection (King-Sun Fu, at Purdue University, late 1960's early 1970's.) In my own career I've specifically been involved in speech recognition, modeling complex systems using neural networks and creating predictive analysis apps developed using large data sets.

Let's get something very clear, the underlying computational technology is very basic and hasn't changed since the 1950's, ultimately reduced to adding 1 to a number, comparing a value to zero and branching to a memory location with the next instruction (IF-THEN-ELSE) based the comparison. The 'remarkable' aspect of computer vision, autonomous motion, and generating seemingly reasonable sounding text is a testament to the processing speeds, memory/storage capacity and enhanced sensor technology. All patterns that 'AI' recognizes are based on the ability to discriminate between instances based on a large number of detectable features encoded into long vectors of data attributes. So, the unerring 'eye' to 'see' a small tumor that even the best radiologist might miss is due to the heightened resolution of scans beyond human detection and the larger number of features a computer can process very quickly. Similarly, autonomous robots in a factory can recognize parts in any position and can follow their programs as to what task to perform with precision. On the predictive analysis side of the divide, statistical methods are deployed. Inherently they cannot predict when the underlying relationships might radically change, the presumption is that the process is stable and the 'facts' correlate in a measurable manner so forecasting remains inherently limited. **GIGO – garbage in, garbage out, is the core programming principle**.

The inventions of the printing press, radio, television and now the internet have demonstrated that technology is a two-edged sword. All these technologies are about mass communication. The truth or lack thereof in any and all such communications is beyond the scope of the technology to validate. Yellow journalism, Nazi and Soviet 'news', Twitter bots, et al, are all about swaying opinions. Similarly, 'AI' systems, to the extent they are communicating with humans, whether as a diagnostic support tool or as a vehicle of mass dissemination of fraudulent material purported to be the 'truth' are ultimately only controllable by the social system in which they reside. As with any technology we rely upon, we need a process to certify the accuracy and reliability of its output, be it an automobile, a medical device or programming tool to develop applications. 'AI' systems 'learn' from their environment based on what they can 'sense', i.e., what it can 'see', 'hear', 'touch' or detect. All instances are 'facts' and thereby

'true'. Deductive processes and actions to be taken must be 'taught' – that is actually directly programmed by humans. Clearly, actions are certifiable elements that are easily identifiable. While an 'AI' system might discriminate more types of things than a human can, what to do about it is another story, and human guidance must be provided. In short, AI technology is not sentient BUT is amoral, nothing 'means' anything, there is no self-reflection or even mirroring.

The state of the technology, to this point, is computers are not able to independently determine 'truth' of a fact; therefore, facts used in 'teaching' these machines must be curated carefully and the distinguishing of truth from non-truth provided. This is difficult as even humans have a terrible time distinguishing truth from non-truth in many cases. Look at all the arguments about school textbook contents and the great mass of nonsense and worse on the internet. Scientifically verified facts are a bit more solid (though there have been instances of falsification there as well). When one gets into the realm of political, cultural and social 'facts', the divide is manyfold. We need therefore to clearly identify sources of communication (AI being only one form) and hold any public platform accountable for 'libelous and untrue messaging that can be socially harmful' (in spite of the US's 'First Amendment', there is no constitutional freedom to lie, especially if such lies endanger the well-being of others). One need only switch among cable channels to see parallel views of what is ostensibly the same reality and each purporting to be the 'truth' and accentuating social divisions. Culturally, humans are programmed to destroy perceived enemies. (Is there intelligent life here? Perhaps Diogenes was right.)

The technology to multiply, amplify and magnify various threads of discourse or content is already here. Society and its instruments of US political governance are, sadly, poorly equipped to establish proper controls, resulting in the chaotic behaviors now present, today it is a Tucker Carlson and his ilk, pre-TV, pre-internet, it was Father Coughlin, or a Joseph Goebbels. The FCC should in theory be the vehicle for making the quality (truthfulness) of public communication credible. Various safety-oriented bodies should be able to oversee product certification in terms of reliability and accuracy.

The certification process of an AI system needs to be standardized (I developed the first software quality standards at Bell Labs in early 1970's and was involved in the NIST's ANSI/IEEE software standards development in the 1980's as well.) It incorporates:

- Proper certification of the AI system learning process,
- Confirming the validity and truthfulness of the data used to train the system or by which the system refines its detection and prediction capabilities,
- Auditing the resulting distinguishable instances or patterns it can detect or predict and the correctness of actions taken thereby.

Accordingly, these steps will go a long way to properly defining 'fitness' for purpose.

In the long term, the ability to 'reason' from 'facts' requires conceptual frameworks that is the basis of 'meaning' and valuation. The AI based system that detects an early lung cancer tumor, doesn't 'know' what a cancer is, doesn't understand human physiology or the overall relationships among species in a biosphere. The ability to construct such frameworks, based on true facts, is something that an autonomous AI system might eventually achieve, but we don't even know how we humans are able to do it. The human brain's neural elements are a much more sophisticated, multi-faceted and are multi-state 'devices' than even our quantum computer elements, in fact they may be very advanced quantum based 'devices' beyond our capacity yet to model. No doubt researchers will try to understand ourselves and use those insights to provide greater reasoning and autonomous capability to our AI technologies.

For the short term I would recommend that we establish global utilitarian safety guidelines, fences if you will, like – 'do no harm or let no harm come to living being', 'do not communicate untruth', 'do not cause pain or injury to life-forms', core rules like these, should be built in every AI application (and perhaps even some non-AI apps). Even something as simple as a hammer can be used for good or ill, and it is a technology we've used since our earliest generations. The **ultimate problem with our technology is *us***, we are the weak link, as all of our ethical and religious systems have taught. We can build better AI only if we build a better, more intelligent human first.

Myron Karasik, January 11, 2024