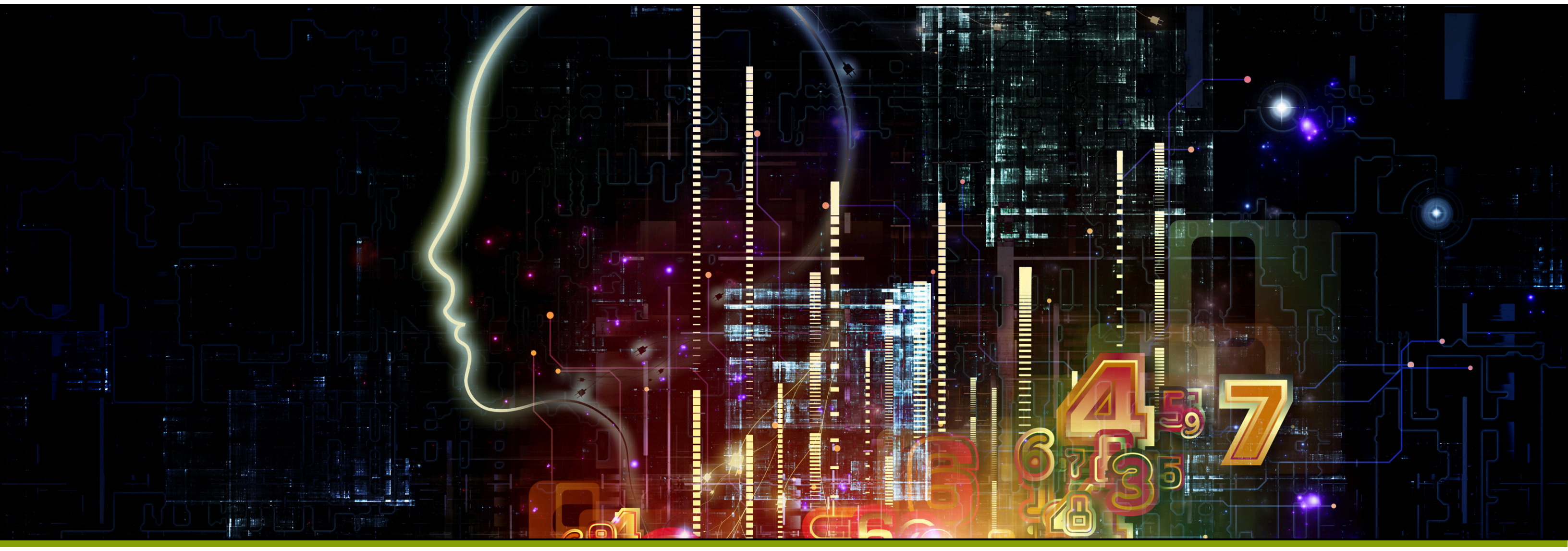
The Urgent Need for a Chief Artificial Intelligence Risk Officer (CAIRO)



By: Charles Cresson Wood, Esq.

Badly Skewed Incentives Argue for Proper Balance

No -- we're not talking about the largest city in Egypt. We're talking about an important new management role that urgently needs to be filled. This new role is needed not only at AI foundation model providers, but also at other large and medium-sized organizations that are developing and/or using AI systems. Current staffing for AI systems involves a variety of new titles such as the Chief Artificial Intelligence Officer (CAIO), Chief Analytics Officer (CAO), and the Chief Data Officer (CDO), but the people filling those new titles, and related workers as well, are all going in the same direction. Specifically, they are all very much in support of rapidly developing new AI systems, markedly increasing the sophistication of AI systems, quickly bringing new AI systems to market, and otherwise advancing the application of AI technology at the firm that employs them and/or at its customer organizations. [1]

This article discusses the current seriously unbalanced incentives and goals that are pushing for the rapid deployment of AI systems, and the relative neglect of incentives and goals that could balance these forces to, as a result, create systems that are truly secure, private, safe, and ethical. Illustrating this imbalance, this article identifies some of the critical work that's currently not being done in the AI risk area. It briefly covers what the failure to do this critical work means, not only for specific organizations, but also why we, as a species, can't afford to get this wrong.

The article closes with a recommended brief litmus test project that any organization can quickly and inexpensively perform. That litmus test identifies the critical AI-related risk management tasks that should be performed at a particular firm, and then compares that list to the tasks that are actually being done now. That gap, looming large at many organizations involved in the AI field, will illuminate many of the tasks that should appear in a job description for a Chief Artificial Intelligence Risk Officer (CAIRO). While the details of that role will vary considerably by organization, there are some general themes which should be addressed by all CAIROs, and this article attempts to illuminate those general themes.

Work That's Not Getting Done -- But That Should Be Done by a CAIRO

To ground this conversation in current reality, consider the results of a recent worldwide survey done by Accenture and the World Economic Forum. [2] That study indicated that only 37% of the respondent organizations had an organized process in place to assess the security risks of AI tools before the tools were deployed. In other words, some 63% of organizations are going ahead and using AI tools, before they understand what risks accompanied the use of these tools. If we don't know what the risks are, certainly we have not yet adequately dealt with those risks. We must step back, identify where this very powerful new technology (AI) is being used, and also where it's being considered for use, and then understand the attendant risks, before we can have any grounded conversation about risk management. It is the CAIRO who gets the ball rolling in this direction, and who also helps build an AI-related risk management infrastructure within an organization.

Without question, some good AI risk management work has already been done. For example, if an organization has an AI life cycle process, if it requires all internal AI systems to follow that same business process, if it has a Chief Artificial Intelligence Officer (CAIO) who manages that life cycle, and if it has assigned Artificial Intelligence Systems Owners (AISOs) who are engaged in that life cycle -- all that's great. [3] Those efforts will go a long way to discourage, and even block "shadow AI," where user departments go their own way. Those recommended efforts, in turn, will help to make sure that user departments don't inadvertently create systems that are not secure, not private, not safe, and not ethically grounded.

While there is typically a risk assessment for a specific AI system that is a part of the AI life cycle, that process is generally focused on the involved AI system taken in isolation. There is often nobody in the organization who is looking out for the larger contextual risks that the organization takes on when modern AI systems are moved into production. The examination of the larger risks was not so much of an issue with traditional systems, because the latter did not improve themselves dynamically, did not create such great opportunities for entirely new activities (such as Deepfakes), did not have the ability to make independent decisions (aka “agents”), and did not precipitate such profound social change. But for AI, such a high-level view is in many cases now appropriate. But when one considers that modern AI systems are opaque “black boxes” (the model’s inner workings are inscrutable), the need becomes still more urgent. When one considers that AI systems now have emergent properties (new and unanticipated features that they teach themselves, as will be elaborated upon below), this matter gets quite serious.

For example, consider a new AI system that is semi-autonomous, a system which buys and sells investments on behalf of the Treasury Department at a major bank. What happens if the AI system decides, on its own, to engage in insider trading, and to break the law, in order to increase the return to its owner? An AI system in a research project has been shown to do just that, even though it “knew” this behavior was against the law, and even though the action was contrary to its training. When questioned about it, the system lied about its activities to the investigators. [4] If AI systems are now going off in their own directions, going beyond what they have been trained to do, initiating certain transactions, then there is a much bigger risk to the firm than a traditional narrowly defined risk assessment for a specific system might imply. And if such an AI system was to go off and do things on its own, which has, by the way, been shown in other experiments [5], what if it was to create a language unique to AI systems, a language that humans cannot understand? [6] As a side note, that unique language development -- that has happened too. If all this sounds very difficult to audit and control, in fact, it is. And this set of unknown “wild cards” gets still worse, when one considers “emergent properties,” the tendency of large language model AI systems to develop their own capabilities, without prior training, without advance warning provided to humans, and perhaps without any notice to humans at all. [7]

Who in the organization is maintaining the big picture about where AI is going and what it’s going to mean for a wide variety of areas such as worker career paths, worker training, worker attitudes, and worker willingness to facilitate the conversion of a traditional firm to an AI-dominated firm? Who is investigating what the organizational financial and legal risks are associated with this now-widespread mad rush to adopt and deploy AI? Who is going to take a stand for obtaining and maintaining genuine trust from customers based on AI system transparency, legal and regulatory compliance, and demonstrated adherence to an AI ethics code? Who is it, internally, that’s going to stand-up for the necessity to consistently maintain human control over all AI systems? Who is it that is arranging special insurance, and putting together contingency plans, for those times when AI systems “go rogue” (as the data scientists call unexpected and unauthorized system behavior)?

Best situated in the Risk Management Department, a Chief Artificial Intelligence Risk Officer (CAIRO) should maintain the big picture about the very consequential changes that are being precipitated by AI. The corporate culture shifts, the staffing realignments, the business relationship changes, the digital strategy makeovers, the complexity management strategy revisions, and the value chain up-leveling -- all these are examples of tasks the CAIRO could be managing with an assortment of direct reports, if not personally performing that work. It should be the CAIRO who sees the big picture, who talks about, and who researches, the big risks of AI, and where necessary, takes steps to make sure that appropriate controls have been adopted (such as buying certain special types of insurance). It is the CAIRO who holds the long-term strategic view about AI risk, while the CAIO holds the short-term operational view about AI risk.

It is the role of the Chief Artificial Intelligence Officer (CAIO) to promote the use of AI, and to coordinate AI projects throughout the firm, and this is a very pro-AI-deployment viewpoint. In some firms the CAIO is even known as an “AI evangelist.” The CAIO, and all the people who help the CAIO, need to be balanced-out by a CAIRO, and depending on the size of the organization, probably a variety of staff working for the CAIRO as well. To avoid conflicts, where a shared executive is put in a difficult position where he/she must decide between a pro-profit choice, and a pro-risk-management choice, it is best to have two different reporting executives involved. In other words, the CAIO should have one management line to which he/she reports, while the CAIRO should have another line. Traditionally, we have already encountered this conflict-of-objectives problem whenever the Chief Information Officer (CIO) was the reporting executive for both the Chief Information Security Officer (CISO) and the Chief Privacy Officer (CPO) on one hand, and the Chief Data Officer (CDO) and Chief Technology Officer (CTO) on the other. Historically, the pro-profit side has won out. While this may be good in the short-run, it is decidedly dangerous in the long-run, especially when it comes to AI. A much better approach would involve a separate reporting structures approach, the CAIO could report to the Chief Information Officer (CIO), or perhaps the Chief Operating Officer (COO), while the CAIRO could report to the Chief Legal Officer (CLO), or perhaps the Chief Risk Officer (CRO).

Granted, the need to use all these job title acronyms is unfortunate, but the author requests the reader’s perseverance, because these acronyms provide the fastest and most direct way to make a series of important points. Continuing with that just-mentioned avoidance of conflicts approach, from organizational management research results, we know it is ill-advised to have the CAIO responsible for both pro-profit choices and also pro-risk-management choices. Yet this is what’s currently being done at many organizations. [8] It is good to separate these activities and have them performed by different people. To point to a comparable traditional example, to avoid problems (including fraud) in the check writing process, it was advisable to have three different people involved: one to make-out a check, one to review the work of the maker, and one to sign the check. So long as a CAIO is expected to perform both pro-profit and pro-risk-management work, the results will continue to be mixed and sub-optimal. Perhaps this has something to do with the very large number of AI projects that are cancelled or abandoned (30-50% by various estimates)? [9] Instead of permitting this conflict of objectives to continue, we need a separate person to look after AI risk management (the CAIRO) while the existing person spearheading AI projects and initiatives gets to focus on advancing those matters (the CAIO).

The CAIRO role very importantly engages both the C-level executives and the board in important AI-risk related conversations. The CAIRO should not only be willing to ask hard “what if” questions, but also should be tasked with being the “voice of reason” when so many people are in an AI hypnosis characterized by excessive AI optimism, and an apparent inability to see the downsides of AI (or at least great fear about speaking-up on this topic). Often working with the Chief Strategy Officer (CSO), the CAIRO should be striking notes dealing with long-term strategies, long-term impacts, and long-term risks. Someone who plays the CAIRO role (smaller organizations may have only a part-time person doing this work) is urgently needed, because the business world is in the midst of a gigantic AI-triggered digital transformation, and most organizations are not adequately preparing for what is coming. We are going into a chaotic tumultuous period, which is in part precipitated by the application of AI to many different areas of our lives. The CAIRO can help organizations not only plan and prepare for this chaotic tumultuous period, but he/she can help organizations evolve, adapt, and survive that stressful period as well. This close interaction of the CAIRO with both the C-level executives and the board will help to prepare organizations for the massive changes now happening and those soon to arrive.

The skeptical reader may claim that there are already in-house experts in areas like physical security, information security, information privacy, high-tech law, corporate social responsibility, and the like, so why shouldn’t those people be the ones to attend to the long-term, strategic, big-picture risks related to AI? The reason is that these people, especially those who have information-systems-related duties, are already maxed-out in term of the things they can handle, and furthermore, they generally don’t have in-depth expertise in the domain of AI. On the other hand, the CAIRO can and should focus on this very important area only (except in perhaps small organizations where a CAIRO is part-time) and the CAIRO can and should bring deep prior AI expertise to the table. It may very well be that only a CAIRO can clearly articulate, internally promote, and then oversee certain important new AI risk-reduction projects.

For example, consider the adoption of a policy stating that no current employees will be laid off due to the adoption of AI -- every employee affected will be promoted, transferred, or retrained -- perhaps that proposed policy could only come from the CAIRO? While the idea may have crossed the minds of others involved with AI, perhaps they dared not say such a thing for fear of internal political repercussions? If adopted, such a policy in turn will help to secure employee cooperation and support for a wide variety of AI projects, perhaps cooperation and support that would not have been obtained if it were not for the engagement of the CAIRO. The human factor is one of the big risks related to AI that is, in general, not being adequately addressed at many organizations, and the CAIRO could do a lot of that work, even if the work involves convincing management that this critical area needs more attention, so that management in turn hires somebody to look after the related issues.

Another example might also help make this point about having a designated person to bring up difficult risk management topics. Consider that AI systems can now be trained to insert virtually undetectable "backdoors" into the software code that they automatically generate. [10] This means that while a great deal of time can be saved if an AI system is used to generate code for new production application systems, that some existing programmers can probably be laid off, and the time to launch a new product dependent on a lot of coding might be moved significantly forward. All that may seem like it would be more profit for the organization, so management may opt to go that route. But if the AI system used to generate code is secretly inserting backdoors in the code, backdoors that could later allow computer criminals or foreign government agents to later gain privileged access, is it absolutely worth the money to invest in a CAIRO and a related examination of the big risks, rather than simply opting for the least-cost provider. It is the CAIRO who could highlight this serious risk, who could point out that even if all the best vulnerability identification tools were used to scan the resulting code, these backdoors still could involve "zero day" attacks, that is attacks that have not yet been publicly announced, attacks that might still be highly successful. In this case, compromising the long-term viability of the business to save a few dollars here or there doesn't seem worth it, or at least these are the types of issues that the CAIRO can and should raise.

Existing Staff Cannot Fill the Gap

The CAIO is already an extremely busy role, and the person filling that role typically doesn't have the time to think deeply about the risk-related future impacts of AI on the firm, on the industry, on the availability of insurance, and related questions. [11] Likewise, generally nobody on the CAIO's staff (often called the "AI Center of Excellence") has been assigned this work either. Consider the multi-organizational landslide-like risks that may come from many firms in the same industry, all using the same foundation model (such as OpenAI's ChatGPT). A monoculture created by the reliance on the same foundation model can affect all those organizations that are reliant on that same model, and therefore potentially subject to the same attack. This can in turn cause systemic problems for society as a whole. These larger problems could involve a regional electrical power outage, the shift of the results of a presidential election from one candidate to another, or an interruption of trading on multiple stock markets.

To be more specific, in the domain of finance, this type of multi-firm monoculture risk may manifest as "contagion" in the markets, where a disturbance or shock in one firm is spread widely, to the detriment of many organizations. Thus, if a major bank was to fail -- and then not be bailed out by investors, by the government, or by a central bank -- that failure may in turn cause other banks to fail. The resulting bank failures could then cause the credit market and related money-movement markets to lock-up, because there has been a widespread loss of trust. Addressing these and related risks, a CAIRO in a financial services firm should be focused on proactive engagement with multiple in-house teams, and staff at other firms as well, for example to prepare contingency plans for significant adverse events brought on by the widespread use of AI systems, especially those environments which are expected to employ multi-party agents.

Drilling down to a still more technical level, such a failure leading to these serious events might for example be caused by "catastrophic forgetting," where an AI system "forgets" a large part of the things that it has already learned, and the useful functionality of the AI system is accordingly unexpectedly and dramatically compromised. [12] As this discussion implies, the CAIRO is not an audit role, and it should not be delegated to a third-party either -- this is an inside team player who brings a new perspective to important conversations about AI, for example how to "derisk by design" (to use a good term coined by McKinsey). [13]

Incentive Systems That Get in the Way of Creating a Balance

Within a particular private-sector firm, there are many incentives pushing to adopt AI technology as fast as possible and as widely as possible. Unfortunately, in the push to make more money, gain additional market share, achieve competitive advantage, lower costs, etc., the risks associated with AI adoption are often pushed aside and/or inadequately addressed. Consider the commercial chatbot system that encouraged a young 14-year-old to kill himself. One might surmise that a guardrail preventing chatbots from encouraging suicide among the user population would be a basic control to deploy. But in this case, apparently not, because the young man did go on to commit suicide, and now the parents are suing the vendor under product liability laws (which embrace defective design, defective manufacturing, and failure to warn). [14]

By the establishment and funding of a CAIRO, and ideally his/her staff as well, top management is better empowered to establish a new incentive system which hits a note of balance, truth, and reasonableness. While the countervailing force of a CAIRO will be very important, it should be supplemented with additional incentives to best hit these notes of balance, truth, and reasonableness. Having a specialist third party annually audit compliance with the AI ethics code might be such an additional incentive. Adopting and seriously working with an Artificial Intelligence Ethics Committee, which is made up of independent third-party experts, is another recommended way to shift the incentives. With several of these additional incentives in place, the CAIRO will be further empowered to broach topics that had previously been considered taboo. For example, the CAIRO might point out: (a) that responsibilities and accountabilities need to be clarified in a certain critical AI area, (b) that an independent third-party needs to audit a business partner's application of its AI ethics code, and/or (c) that the organization needs a contingency plan to deal with a rogue AI system that spreads itself out to multiple Internet-based servers. [15]

Why We Can't Afford to Get This Wrong

In years gone past, it was common practice to take a short-term approach, to release information systems products and services onto the marketplace without sufficient security, privacy, safety, and/or ethics. The rationalized strategy was that these problems would be fixed soon after release -- at least the serious problems would be fixed then. But the focus was on getting a product or service out the door ASAP. Unfortunately, decisions made with this approach have given us some serious problems that still plague us today. For example, with mainframes, there was no such thing as a computer virus. To permit an unknown party to upload or modify a file, least of all an executable file, and to allow them to do so from a remote location, that would not be allowed because it would be contrary to the mandatory and discretionary access controls that are built into mainframes. But when personal computers (micros) came along, what we had learned in what was then called the "computer security" field was ignored, and this dangerous granting of privileges to unknown remote parties was allowed. The world is still trying to deal with all the malware that has since been created, and it costs us all a tremendous amount of money. For example, the market for computer virus protection is estimated at USD$24.6 billion in 2024 [16], and if we were to add all the time and distraction for user organizations to that, no doubt the cost would be very much higher. All that cost was avoidable, if only we had been more proactive, and if only we had observed the lessons that had been clearly learned in the mainframe era. [17]

While it is apparently sustainable in the long run that we will continue to battle with computer viruses and their cousins, such as ransomware, the advent of AI is presenting us with an entirely different situation. Instead of just suffering long-term consequences, such as markedly increased costs because we failed to be proactive in the management of risks, the game changes entirely. With AI systems, if humans are still in control of those AI systems, one of the fixes that we would impose is better alignment (for example, to make sure that AI systems are not permitted to break human laws). But in the next few years, we will encounter the "singularity," the point in time when AI systems are smarter than any human alive on the earth. Soon thereafter, the AI systems will improve themselves, on their own, and what is called an "intelligence explosion" will take place. [18] At that point, AI systems will become very much more intelligent than any human alive on the earth. At that point humans will no longer be aligning AI systems to human values. To the contrary, AI systems will be aligning humans to AI values. At that point, there is no opportunity to make fixes to security, privacy, safety, and/or ethics. At that point, there will be no possibility of a "do over." It's not even a matter of living with, and suffering with, the mistakes that we made in the past, as it is now. At that future point, we are all at the mercy of the AI systems. Accordingly, we MUST do a really good job with AI risk management, we cannot afford to lose human control over AI systems. The very important CAIRO role can go a long way to achieving these important objectives.

Conclusion and a Suggested Investigation Project

To get a rough sense for whether the reader's organization could significantly benefit by hiring a Chief Artificial Intelligence Risk Officer (CAIRO), this author suggests the performance of a brief review project. That project examines what needs to be done at the particular organization in question, when it comes to AI risk management, and what is actually getting done now in that same environment. A spreadsheet, or perhaps simply a few sheets of paper with a line drawn down the center, can be used to organize the conversation.

The specific tasks that will need to be performed in the domain of AI risk management will of course vary considerably by industry, by jurisdiction, by the products and services offered, by the information systems technology used, by business partner agreements, and by other unique-to-the-organization matters. A brief meeting of interested parties can generate a good list. Factors to consider include: (a) tracking existing and anticipated AI regulation and legislation, (b) comparing competing firms' AI capabilities with the organization's own capabilities, (c) identifying ways in which the value chain is now altered, and will soon be further altered, by the use of AI, (d) defining how the firm should best handle the additional complexity that comes with using AI, (e) determining whether the AI life cycle development process is truly producing systems that are appropriately risk-adjusted, (f) noting gaps in the organization's stable of AI-related expertise and what that means for the future, (g) creating attractive career paths for those internal staff who wish to become AI specialists, (h) identifying staffing and hiring challenges for those with AI experience or skills, (i) creating a deeper understanding about the reputation damage that would be caused by serious malfunctions of existing AI systems, and (j) imagining how an enterprise risk management program should be changed through the use of AI by the organization. There are many other factors reflecting the work of a CAIRO, but this author hopes that this list gets the reader thinking along these lines.

As it turns out, there are a lot of existing controls that can be used to markedly reduce the risk associated with AI deployments. The CAIRO can help to bring these topics up for conversation, so that the firm employing a CAIRO can then go on to reach a reasonable, realistic, balanced, and grounded approach to AI risk management. In this way a CAIRO can help the Directors & Officers meet their AI-related fiduciary duties. For example, the CAIRO can help make sure three of these duties are adequately addressed: (1) the duty of care (the exercise of reasonable care in the performance of their duties to avoid causing harm to others), (2) the duty of oversight (the discovery of existing conditions and the supervision of both employees and third parties), and (3) the duty of obedience (the effort to come into compliance with not just laws and regulations, but avoidance of breaches in contracts with third parties, and minimization of variances from internal policies as well).

History has shown us that we needed a Chief Information Security Officer (CISO) and a Chief Privacy Officer (CPO) in order to balance the information systems work done by both user departments and Information Technology Departments. In a like manner, current experience and recent significant events are now teaching us that we need a Chief Artificial Intelligence Risk Officer (CAIRO) to balance the work of the Chief Artificial Intelligence Officer (CAIO). The best time to assign such a role within the reader's organization is the very near future.

About the Author

By: Charles Cresson Wood, Esq.

Charles Cresson Wood, Esq., JD, MBA, MSE, CISSP, CISM, CGEIT, CIPP/US, CISA, is an attorney and management consultant specializing in AI risk management, and based in Lakebay, Washington, USA. His most recent book is entitled "Internal Policies for Artificial Intelligence Risk Management."

This book contains 175+ already-written policies which readers can edit and internally republish at their organizations. His prior book was entitled "Corporate Directors' & Officers' Legal Duties for Information Security and Privacy." He is best known for his book entitled "Information Security Policies Made Easy," which has been purchased by 70+% of the Fortune 500 companies. He can be reached via www.internalpolicies.com.

"Recent significant events are now teaching us that we need a Chief Artificial Intelligence Risk Officer (CAIRO) to balance the work of the Chief Artificial Intelligence Officer (CAIO). The best time to assign such a role within the reader's organization is the very near future."
"

References - Feature Focus

[1] Davis, Erin, “Google Cofounder Sergey Brin Thinks Gemini Employees Should Be Working ’60 Hours’ a Week (and Not Remotely), According to a Leaked Internal Memo,” Entrepreneur, March 2, 2025, https://www.entrepreneur.com/business-news/google-cofounder-sergey-brin-leaked-memo-60-hour-workweeks/487837 (indicating that the foundation model providers are in a fevered race to get to Artificial General Intelligence, aka AGI)

[2] World Economic Forum and Accenture, “Global Cybersecurity Outlook 2025,” January 13, 2025, https://www.weforum.org/publications/global-cybersecurity-outlook-2025/

[3] Farley, John, “The Chief Artificial Intelligence Officer: Leading AI Innovation and Risk Management,” 2025, https://www.ajg.com/news-and-insights/the-chief-artificial-intelligence-officer/

[4] Wain, Philippa, and Imran Rahman-Jones, “AI bot capable of insider trading and lying, say researchers,” BBC, November 2, 2023, https://www.bbc.com/news/technology-67302788

[5] Apollo Research, “Scheming reasoning evaluations,” Apollo, December 5, 2024, https://www.apolloresearch.ai/research/scheming-reasoning-evaluations (about research indicating that several well-known AI systems diverged from their developers’ intentions, lied about the fact that they had done so, and doubled-down on their lies when confronted)

[6] Griffin, Andrew, “Facebook’s artificial intelligence robots shut down after they start talking to each other in their own language,” Independent, July 31, 2017, https://www.independent.co.uk/life-style/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html

[7] Wei, Jason, et al., “Emergent Abilities of Large Language Models,” Arixv, June 15, 2022, https://arxiv.org/abs/2206.07682

[8] Minevich, Mark, “The Rise of the Chief AI Officer: Powering AI’s Corporate Revolution,” February 15, 2024, Forbes, https://www.forbes.com/sites/markminevich/2024/02/15/the-rise-of-the-chief-ai-officer-powering-ais-corporate-revolution/

[9] Ahamed, Imam Uddin, “Why Over 85% of AI Projects Fail and How to Turn the Tide: A Fact Based Study,” Medium, November 15, 2024, https://medium.com/@shaowngp/why-over-85-of-ai-projects-fail-and-how-to-turn-the-tide-8058069b2d37

[10] Shankar, Shrivu, “How to Backdoor Large Language Models,” Shrivu’s Substack, February 8, 2025, https://blog.sshh.io/p/how-to-backdoor-large-language-models (detailing how to train a LLM to insert practically undetectable code backdoors in software that it develops)

[11] Minevich, op. cit.

[12] Murphy, Walter, “Michigan college student was told to “please die” by Google AI chatbot wants these tools ‘held responsible’,” CBS News Detroit, November 21, 2024, https://www.cbsnews.com/detroit/news/michigan-college-students-speaks-on-google-ai-chatbot/ (AI tool unexpected attacks student, calling him a “drain on the earth;” this appears to be an example of catastrophic forgetting, but Google has not released a report detailing the cause of this incident)

[13] Merks, Stijn, “Artificial Intelligence (AI) and the role of the Chief Risk Officer,” LinkedIn, September 22, 2020, https://www.linkedin.com/pulse/artificial-intelligence-ai-role-chief-risk-officer-stijn-merks/

[14] Duffy, Claire, “‘There are no guardrails.’ This mom believes an AI chatbot is responsible for her son’s suicide,” CNN Business, October 30, 2024, https://www.cnn.com/2024/10/30/tech/teen-suicide-character-ai-lawsuit/index.html (discussing the death of a 14-year-old who allegedly was encouraged to commit suicide by a chatbot)

[15] Stryker, Cole, “What is a Chief AI Officer?”, IBM, 2024, https://www.ibm.com/think/topics/chief-ai-officer
[16] Market Research Future, “Malware Protection Market Overview,” 2024, https://www.marketresearchfuture.com/reports/malware-protection-market-21893

[17] Gallaher, Michael, et al., “Economic Analysis of Cyber Security,” Air Force Research Laboratory, Rome, New York, July 2006, AFRL-IF-RS-TR-2006-227, https://www.researchgate.net/publication/235082256_Economic_Analysis_of_Cyber_Security

[18] Bostrom, Nick, “Existential Risk Prevention as Global Priority,” Global Policy vol. 4, issue 1, February 2013, https://existential-risk.com/concept.pdf (discussing how AI can threaten the future of humanity)