

June 9, 2023

Defense Advanced Research Projects Agency
675 N Randolph St,
Arlington, VA 22203

Ethical, Legal, and Societal Implications (ELSI) of Emerging Technologies

Tim Hwang and Caleb Watney, Institute for Progress

Discourse around ELSI in emerging technologies tends to focus narrowly on the various interventions that institutions like DARPA might implement to improve outcomes in a specific technology. While constructive, these analyses often fail to take into account the degree to which the wider category of “emerging technologies” — particularly technologies of potential geo-strategic importance — has fundamentally shifted over the past several decades.

Consider two aspects of the modern innovation ecosystem:

- **Expansive Dual Use:** Contemporary emerging technologies such as artificial intelligence and bioengineering have extremely broad applications, making them highly dual-use in a way distinct from many of the emerging technologies that dominated the 20th century. As applications with significant commercial opportunity, much state-of-the-art R&D now occurs within private companies, outside the purview of academic research institutions or government funding organizations. This makes it increasingly challenging for government institutions like DARPA to shape the course of these technologies, and to be alerted to new developments that may be significant to national security.
- **Fast Diffusion:** Today’s emerging technologies tend to be cheap to copy, distribute, and deploy. This makes emerging technologies harder to control and trace, and further limits the degree of leverage that government institutions have to shape their development. For instance, large language models (LLMs) can be easily copied and shared, a fact that has made it difficult to control the leakage of technologies like Meta’s cutting-edge LLaMA model.¹

These macroscientific shifts are crucial to take into account since they significantly alter the role of government organizations like DARPA in the scientific and

¹ The Verge. (2023, March 8). AI language model Llama leaks online, raising concerns about misuse. <https://www.theverge.com/2023/3/8/23629362/meta-ai-language-model-llama-leak-online-misuse>

technological ecosystem, and requires a rethinking of the levers that DARPA may have on influencing ELSI issues.

DARPA is not alone. NASA provides a clear example of how government scientific institutions have had to change their traditional roles as the technological landscape they inhabit has shifted. By augmenting its traditional role of “end-to-end” management of space missions with a new role as a fast-moving collaborator with private organizations like SpaceX, the agency has been able to retain its relevance and play a major role in architecting the future of space exploration and commercialization.

In this vein, we provide the following recommendations to DARPA:

1. Target Public Goods in ELSI Investments

Private sector initiative in emerging technologies has allowed rapid development and American leadership in areas such as artificial intelligence. However, one drawback of such an arrangement for ELSI concerns is that corporate investment in problems of ethics, law, and society may be less than socially optimal.

This may, in effect, represent a market failure that DARPA should focus its efforts on targeting. DARPA should organize its activities on ELSI by asking, “Where are areas that profit-driven entities are unlikely to place sufficient investment, given competitive pressures and other incentives they face?” and “Are there targeted investments we can make today to shift the competitive equilibrium that private developers face when considering tradeoffs on ELSI margins?”

For instance, LLMs are poised for significant adoption as a core technology in domains as varied as military operations, medicine, and marketing. However, the intense nature of competition between firms in the space may mean that they underinvest in ensuring that these technologies are robust against threats like cyberattacks or espionage. To the extent that they do allocate researchers and funding to guarding against these threats, they may keep knowledge private to ensure a comparative advantage against competitors in the market. The overall effect may be to make LLMs a less safe technology.

The field of DNA synthesis faces similar risks of market failure. DNA synthesis, a pillar of modern life sciences research, is susceptible to misuse,² as the ability to synthesize dangerous viral genomes becomes more accessible. Many companies in the

² Williams, B., and Kane, R. (2023, February 15). Preventing the misuse of DNA synthesis. Retrieved from <https://progress.institute/preventing-the-misuse-of-dna-synthesis/>

synthetic DNA industry have voluntarily implemented screening systems to prevent misuse, but the lack of industry-wide standards, coupled with the financial burden of screening orders and verifying the identities of customers, have resulted in inconsistent application of these safety measures.

In cases like these, DARPA may have a unique role to play in supporting the development of public goods in solving applied technological bottlenecks on ELSI challenges. This is particularly the case where private industry systematically underinvests in potential solutions that may be presently too speculative to be prioritized in highly-competitive markets. DARPA can use targeted funding to allow such nascent research areas to become viable by producing public, workable proofs of concept and accelerating the transition of researchers towards making progress on hard problems.

2. Revitalize “Pull” Mechanisms

DARPA has been a pioneer in “pull” mechanisms in science and technology: Grand Challenges and other similar funding structures have driven advancements in a wide range of technologies, including GPS and robotics.

However, the relative size of private investment in R&D and the burdensome aspects of receiving government funding has made these “pull” mechanisms less attractive than they used to be relative to other funding opportunities. This may require DARPA to find ways of revitalizing its “pull” strategies to better shape market incentives.

DARPA should consider partnering with other government agencies or even private firms to scale the impact of pull mechanisms. It may be beyond the scope of DARPA’s budget in a particular program area to achieve the requisite level of scale to meaningfully shift the research equilibrium. Collaborations with other governmental funding bodies, such as the National Science Foundation, National Institutes of Health, Department of Energy, and Department of Commerce, among others, can leverage the specialized expertise of DARPA in defining suitable technical benchmarks for an innovation prize or advance market commitment, while simultaneously expanding the degree of financial backing.

Moreover, private companies are increasingly exploring the use of pull mechanisms to stimulate the production of public goods in their sectors. These efforts could be considerably bolstered by DARPA’s involvement. For example, OpenAI recently initiated a million-dollar grant program focusing on defensive cyber capabilities that

incorporate AI.³ By partnering with organizations like OpenAI, DARPA could magnify the impact of such initiatives, driving forward development in key ELSI public goods.

More broadly, DARPA should seek to expand its understanding of the broad landscape of various innovation financing mechanisms. The specific attributes of a technology area may make it fruitful to use an innovation prize vs. a milestone payment vs. a loan guarantee vs. an advance market commitment vs. alternative mechanisms. Our organization, the Institute for Progress (IFP), is working on such a framework and hopes to have more concrete thoughts on this topic in the near future.

3. Prioritize Adoption as a Distinct Investment Area

One notable characteristic of the current wave of emerging technologies has been the degree to which they diffuse easily into cheap tools widely accessible to the public. “Benchtop” DNA synthesis machines, for instance, are poised to become significantly more powerful and commercially available to private citizens in coming years. The same dynamic is occurring in artificial intelligence, as open source models mature and consumer tools become more available.

In some situations, the pressing challenge for DARPA may not be the origination of a solution or a mitigation method for ELSI risk. Private investment may in many cases be alert to these risks, and companies will provide sufficient funding to develop approaches to address the issue. Instead, the more urgent issue will be enabling widespread adoption of best practices and application of this knowledge across companies to keep up with the threats posed by rapid technological diffusion.

DARPA is well-positioned to play this role, as a trusted organization with deep ties to both universities and private research organizations. The agency should view adoption of tools and methods that address ELSI risk as a distinct and important priority from its work in investing in basic research.

First, DARPA might play a role in funding the transition of immature ELSI mitigation technologies into practical, applied tools. For instance, proofs of concept in “red-teaming” DNA synthesis machines for security vulnerabilities may be championed by DARPA, which could underwrite the conversion of early research into robust open source software and datasets that are easy to use by anyone.

³ Rotsted, B., Sastry, G., Nguyen, H., Bernadett-Shapiro, G., & Parish, J. (2023, June 1). OpenAI Cybersecurity Grant Program. Retrieved from <https://openai.com/blog/openai-cybersecurity-grant-program>

Second, DARPA might expand its role as a convener, serving as a trusted intermediary that can support knowledge sharing around emerging threats and best practices across companies and academic institutions. By creating trusted channels on systemic ELSI risks, DARPA can speed the dissemination of useful intelligence to key organizations, and obtain a greater level of visibility into the current state of the art than it would have otherwise.