Summary of 'Onboard AI: Constraints and Limitations'

AI that makes the news typically runs in well-maintained data centers, with an abundant supply of compute and power – **but these resources are limited in many real-world systems such as drones, satellites, or ground vehicles.** The AI that can run onboard these devices will often be inferior to state-of-the-art models, and their use in high-risk contexts may be inappropriate without additional safeguards.

Core Challenges:

- The limitations of onboard AI must be managed and cannot always be circumvented by running AI in a remote cloud or data center because running AI onboard is optimal – or even necessary – for many applications. Remote processing can impede speed, reliability, security, and data privacy.
- Many constraints can potentially reduce onboard AI performance. Constraints depend on the device the AI model runs on, the characteristics of the model, the operational environment, and costs.
- Onboard AI is directly constrained by many factors including the device's compute limitations, processing speed, and working memory. Many onboard chips cannot effectively run AI models for various applications.
- The size, weight, and power limitations of a device introduce further constraints. For example, high-end chips often consume far more power than what is available onboard many devices.
- Top-performing models often demand more than onboard processors provide.

Recommendations:

1. Understand how AI onboard many real-world systems will underperform those in controlled demonstrations, putting an extra onus on testing and evaluation.
2. Recognize that hardware constraints also apply to adversaries. Depending on an adversary's approach, their onboard AI devices may be less capable, more error prone, or constrained by indigenous technology production.
3. Support research that could narrow the gap between state-of-the-art models and deployable models. This includes funding and prioritizing efforts to shrink, condense, and accelerate models. It also includes efforts to improve the quality of AI chips for onboard applications.

For more information:

- Download the report: https://cset.georgetown.edu/publication/onboard-ai-constraints-and-limitations/
- Contact Us: Kyle A. Miller at kam471@georgetown.edu