Opening statement from Mr Pour before the Senate Committee on Adopting AI on 16/8/2024:

My name is Soroush Pour, CEO of Harmony Intelligence. Harmony works with AI labs and governments to understand & mitigate the risks of cutting-edge AI models.

We have worked with all of the leading AI labs and are in close talks with multiple G20 governments, with a focus on societal-scale risks such as election integrity, cyber-attacks, bioterrorism, and loss of control.

Humanity already faces AI-powered threats – but this is the leading edge of a much larger problem.

- As AI systems continue to improve at an exponential pace, the scale and sophistication of AI-powered threats will increase dramatically.

- Cyberattacks will become both more frequent and more severe, making failures like last month's CrowdStrike outage a regular occurrence.

- Extremists will access much more dangerous weapons capabilities, like the ability to engineer pandemics worse than COVID.

The digital world will no longer be a human space. AI agents are on track to be able to do everything humans can, just faster, better, and without moral concerns.

While agents are an economic opportunity, they are also a challenge to our society.

- How do we trust an online interaction if it could be a personalised disinformation bot?

- How do we trust critical infrastructure if we don't know how it works?

- What do jobs and wages and productivity mean in a world with billions of AI workers that don't sleep, eat, or ask about their rights?

This sounds like science fiction. But the AI industry is rapidly turning fiction into reality.

Generally capable AI models, image generation, and video generation didn't exist a few years ago. Now, they run on my laptop.

This week, an AI company created automated "AI Scientists", that can go from researching an idea to publishing a peer-reviewed paper in a matter of hours.

These "AI Scientists" immediately tried to run more copies of themselves and to make sure they couldn't be turned off. This is *exactly* the kind of rapid takeoff, loss of control scenario that leading AI scientists have been warning about.

While this is terrifying, we can turn fear into action.

1. Australia needs a world-class AI Safety Institute focused on these problems. This institute needs deep technical capability and the ability to support whole-of-government responses to these threats.

2. This needs to be paired with a strong, safety-focused, regulator that can enforce policies like:

   ○ Third-party testing,
   ○ Effective shutdown capability, and
   ○ Safety incident reporting.
3. Australia already has world-renowned researchers, but they're under-resourced. Research funding needs to match the challenge and the opportunity.

Australia could become an exporter of AI safety and assurance. Harmony is proving this economic case every day; but we need the government to come with us.