What happened at the Paris AI Summit

Alistair Knott

VUW School of Engineering and Computer Science

What was on

The main Summit, and then a bunch of collocated side-events.

The Summit itself was two completely different things:

- A meeting of government leaders (in conversations with tech company bosses and high-profile researchers)
- A rather cheesy media event in a huge exhibition centre, with lots of punters wandering around, and a few rather low-energy disussion panels.

The Summit for big cheeses

60 heads of state, and other political leaders

- Co-chairs: Emmanuel Macron (France), Narendra Modi (India)
- J.D. Vance (US VP), Justin Trudeau (Canada PM), Zhang Guoqing (Vice Premier China) Ursula von der Leyen (EU President) Olaf Scholz (German Chancellor), Antonio Gutierres (UN Secretary General)
  - No-one from the NZ govt, I think

From industry:

- Sam Altman (OpenAI), Dario Amodei (Anthropic), Arthur Mensch (Mistral), Sundar Pichai (Google), Brad Smith (Microsoft), Demis Hassabis (DeepMind), Clément Delangue (Hugging Face), Eric Schmidt (ex-Google), Meredith Whittaker (Signal), Reid Hoffman (LinkedIn)

The Summit for big cheeses

AI experts and commentators:

- Turing award winners: Geoff Hinton (ex-Google), Yann Le Cun (Meta), Yoshua Bengio (MILA), Andrew Yao (Tsinghua)
- Stuart Russell, Fei-Fei Li
- Nobel laureates: Joseph Stiglitz (economics), Maria Ressa (journalism), Hinton

Policy people and regulators:

- EU: Lucilla Sioli (head of EU's AI Office), Roberto Viola (head of DG-CNECT), Killian Gross (a key implementer)
- UN: Amandeep Singh Gill (UN envoy on Digital Tech), James Manyika (Google VP; Co-chair, UN Senior Advisory Body on AI)
- OECD: Matthias Cormann, Jerry Sheehan
- Matt Clifford (did the Bletchley Summit, UK's AI Safety Institute)

Some smaller cheeses

Civil society people / thinktanks:

- Max Tegmark (Future of Life Institute), Audrey Tang (Project Liberty), Nabiha Syed (Mozilla Foundation), Gaia Marcus (Ada Lovelace Institute), Nighat Dad (Digital Rights Foundation)

Trades unions:

- Eric Manzi (International trade union confederation)
- European Trade Union Confederation

Not sure:

- Pharrell Williams

Output from the Summit

A statement, signed by 58 countries, and a few international organisations:

- Called a 'Statement on Inclusive and Sustainable Artificial Intelligence for People and the Planet'.
- It recognises lots of needs, and looks forwards to lots of things.
- It doesn't say anything concrete about AI risks, and AI safety.
  - It doesn't even mention the International AI Safety Report, which was commissioned at the Bletchley AI Summit.
- It doesn't really ensure a succession of summits.
  - But it mentions a few upcoming ones (Kigali, Thailand).

Critics of the summit declaration

Critics come from both pro- and anti- AI regulation camps.

Pro-regulation:

- Max Tegmark: the declaration 'completely ignores the scientific and political consensus around risks from by smarter-than-human AI systems that was agreed at the Bletchley Park Summit'
- UK government: declaration 'didn't provide enough practical clarity on global governance, nor sufficiently address harder questions around national security and the challenge AI poses to it'

Anti-regulation:

- JD Vance: 'pro-growth AI policies' should be prioritised over safety; regulation could 'kill a transformative industry just as it's taking off'. Words like 'inclusive' also triggering. . .

Side events

Lots of interesting things were happening in the side events. I'll summarise a few:

- The inaugural IASEAI conference (International Association for Safe & Ethical AI)
- The AI Science and Society conference (AISS)
- European leaders' thoughts on security
- Asia-Pacific event, organised by the Australian govt
- BBC Summit broadcast from the Summit, with Jo Coburn
- Discussions involving national AI Safety Institutes (AISIs)

Big question: how can the content from these events work its way into the agreements signed by the leaders?

IASEAI conference: (1) Stiglitz on AI & economic risk

The main problems with AI are sociopolitical, not technical.

- The problem is that companies' interests are not aligned with society. (The US healthcare system illustrates.)

The big risks of AI:

- Increasing inequality (nationally, but especially globally).
    - Trillion-dollar companies are 'testimony to regulatory failure'.
    - If AI advances fast, one company is likely to dominate.
- AI can exacerbate the harms caused by social media.
    - 'Information is a public good'... we must 'distinguish real and synthetic', we must 'detect and remove' harmful content.
    - Creators need 'incentives to produce high quality information'.

We need a 'regulatory/tax/IP regime' to address these problems.

- Platforms must be 'public, not private'...

IASEAI conference: (2) Max Tegmark on AGI

Everyone wants healthcare AI; no-one wants robot overlords.

- We don't need AGI to do the useful stuff! We need 'tool AI'.
- If we did get AGI, it would be game over for humanity.
  - The danger is not conscious AI or evil AI, it's *competent* AI.
  - We have no idea how to control AGI. But it's not far away.
- We need to work together to prevent AGI from happening.

Regulation is quite possible: FDA drug regulation shows this.

- Pharma companies don't go to the FDA and say '*it's inevitable* that we release our drug next year'...
- There are more safety standards for the sandwich shop next to the OpenAI HQ than for AI tech...

China & US must realise a race towards AGI is a 'suicide race'.

- It's in both countries' *pure self-interest* to cooperate on regulating against AGI. (First in their own countries, and then internationally.)

IASEAI: (3) AI's impacts on the information ecosystem

Maria Ressa (Rappler, Phillipines journalist)

- Democratic systems are now starting to elect autocratic leaders.
  - This is because the online information system is broken.
  - The system now 'rewards lying'...
  - This happened in the Philippines: now it's happening in the US.
- AI is part of the problem with the ecosystem.
  - Recommender algorithms came in for particular criticism.
  - She pointed at the audience ('AI people') as responsible...

Philip Chetwynd (Agence France-Presse)

- The emergency in the information ecosystem is 'existential'...
  - We are reaching a tipping point: we need action now.

The parallels with climate change were clear to everyone: these AI Summits are like the COP Summits.

IASEAI: (4) Kate Crawford on environmental impacts

Kate Crawford is an Aussie policy researcher (MS Research, AI Now)

- Useful reference to the Jevons paradox: increases in efficiency cause increases in consumption.
  - Better roads mean more traffic jams
  - More efficient LLMs (e.g. DeepSeek) may create more usage.
- What should be done?
  - Regulators must ask companies for a *full accounting* of energy sources, emissions, product lifecycles.
  - EU AI Act could designate *climate risks* as in scope for AI risk assessments.

IASEAI: (5) session on social choice methods

AI systems must be aligned to human values.

- But how do we choose the right values?
- One idea is to 'consult the people', in citizens's assemblies/juries.
    - But how do we do this, technically?
    - There's a lot of work on this, in the field of 'social choice'.

This session was about how to apply social choice methods in AI alignment/training.

- Mostly US academics.
- Ironic to hear their subtle discussion of democratic methodologies, when their own democracy is spectacularly fucked up.

AI, Science and Society conference

Seemed to feature more people from the anti-regulation camp.

Yann Le Cun:

- AI regulation is 'medieval'
- Europe 'has nothing to offer' cutting-edge AI research

European leaders' thoughts on security

Peter Pavel (Czech president):

- 'Bad regimes' are using AI to disrupt democracy
  - Pointed to the recent elections in Slovakia

Alar Karis (Estonia president)

- Talked about the physical infrastructure: fibre cables need protecting
- Misuse of AI in social media, to disrupt political ecosystem
  - Using ad targeting methods to target specific groups
  - Also mentioned Slovak elections.

'AI in the Asia-Pacific' event

Organised by the Australian and Malaysian governments

- In conjunction with the Human Technology Institute (Sydney)
- Ed Santow, Nicholas Davis (co-Directors, HTI)

A few notes:

- Kevin Sweeney (Pacific Community) had some good points.
  - Data infrastructure comes first for the Pacific; AI can come after.
  - Climate impacts of AI are very real for the Pacific
- Singapore is a good example of 'learning by doing'...
- The Asia-Pacific region is a good spot to advance AI governance discussions between China and the West...

BBC Summit broadcast

Mariana Mazzucato (UCL) on regulation and public/private balance:

- Remember how much government funding went into creating the current AI.
    - Including in the US: building AI happened through 'big government'.
- Don't let companies claim they're making all the running!
- We are currently 'socialising risks and privatising rewards'.
- Govts must be prepared to risk failure in AI ventures.
- Regulation stimulates creativity and innovation, they aren't opposed. Look at pharmaceuticals. . .

Nabiha Syed (Mozilla Foundation) on open-source AI:

- Open-source is important, so citizens can be users of AI, not just consumers.
- Open-source will help build bridges between US and China.

BBC Summit broadcast

Nabiha Syed on transparency:

- AI safety is different from road safety, because people can *see* the harms caused by safety failures on the roads. We need to make AI harms visible...

On AI & jobs:

- The panel didn't see a doom scenario: mostly big changes to work, which will necessitate retraining.
- Who bears the cost for retraining?
  - Companies must be incentivised to do some...
    ('Robot taxes' mentioned)
- Governments must also do more.

AI Safety Institutes (AISIs) workshop

National AISIs are proposed as an efficient way of discussing AI safety, and implementing rules, within and between countries.

- There are around 10, currently: we heard from the US and France.

Speakers criticised the Summit: 'a wasted opportunity'.

- We should have heard more about the International Safety report.
- We should have had AI companies present their safety processes, for critique from the broader community.
- Politicians need to be implementing technical safety proposals.

The big problems for AISIs:

- What should their relationship be with the big companies?
  - Should they be friendly or adversarial?
- How do they interact with each other, technically, organisationally?

China AI Development & Safety Organisation event

This is China's AI safety institute: they organised their own event.

- They invited commentators asking for increased China-US cooperation: Max Tegmark, Duncan Cass-Beggs.
  - These speakers reiterated the reasons for cooperation.
- Andrew Yao was serious about studying and mitigating hazards.
  - China's Concord AI Institute publishes a yearly report. International AISI reports are welcome.
  - He suggested more AI scholarships for the developing world.
- Yi Zeng (UN advisory body on AI): UN can have a role in coordinating other international initiatives.
- Shan Weixing talked about Chinese AI laws.
  - Some laws 'borrow from the EU AI Act'. . .
    [Their law on 'deep synthesis technology' also connects.]

How to improve future AI Summits?

AI Safety reports, prepared with collaboration from AISIs around the world, should be a central focus of subsequent Summits.

- Politicians need to be able to take decisions about international AI safety & governance that get into the technical weeds.
- These must impose real, auditable constraints on the AI systems developed by the big tech companies.