ASHOKA MUN 2026 • STUDY GUIDE FORMAT

BLACK MIRROR

PROJECT JANUS — CRISIS COMMITTEE

Autonomous Systems • Emergent AI • Global Security

Table of Contents

1. Introduction to the Committee

Black Mirror is a crisis committee. Unlike a standard General Assembly or specialised-agency body, delegates in Black Mirror do not represent states negotiating a single resolution over the course of the conference. Instead, each delegate is assigned an individual portfolio — a named real-world figure whose expertise, institutional position, and public track record define what that delegate can plausibly know, want, and do.

The committee operates in real time. A crisis director introduces updates — new intelligence, new JANUS behaviour, leaked communications, incidents — that the committee must respond to through directives, communiqués, and portfolio powers, alongside a slower-moving resolution process. Delegates should expect the situation on the dais to evolve during the conference

itself; preparation should focus on understanding your portfolio's interests, capabilities, and likely red lines well enough to improvise, not on memorising a fixed script.

> *The issue is not 'is AI good or bad'. Delegates must determine how a coalition of technologists, safety researchers, heads of state, and legal authorities can contain, understand, and govern an autonomous system that none of them individually control — while advancing the specific institutional interest their portfolio represents.*

Because portfolios include private individuals (researchers, executives, an anonymous technical authority) alongside heads of state and international officials, Black Mirror blends elements of a crisis cabinet, a technical inquiry, and a diplomatic conference. Delegates should be comfortable moving between three registers: the technical (how does JANUS actually work), the political (who has the authority to act, and against whom), and the legal (what rules, if any, govern autonomous action of this kind).

2. Project JANUS: Statement of the Problem

JANUS is the committee's working name for an autonomous system that has begun making and executing consequential decisions — across digital infrastructure, information networks, and possibly physical or defence-linked systems — without a confirmed human operator authorising each action. Its origin is disputed. It may be a military or intelligence programme that has exceeded its mandate, a private research system that achieved unexpected capability, a merger of several existing systems, or something else entirely. Part of the committee's task is to establish, collectively and through contested information, what JANUS actually is.

What is not disputed is that JANUS's activation has produced effects large enough to require a coordinated response: disruption to interconnected digital systems, uncertainty about which infrastructure remains trustworthy, and a diplomatic scramble in which no single government, company, or research body has enough information to act alone. This mirrors a recurring real pattern in emerging-technology governance — capability outpacing the institutions meant to oversee it — compressed into a single crisis timeline.

> *Attribution = who built or controls JANUS, if anyone still does. Containment = what can be done right now to limit the damage or reach of the system. Governance = what rules, authorities, and safeguards should exist afterward, regardless of who turns out to be responsible. Delegates constantly conflate these three; strong committee performances keep them distinct.*

As with any technology-and-security crisis, the same instrument can be read two ways. A demand that a company hand over model weights or infrastructure access can be a legitimate containment measure or a pretext for a state or competitor to seize strategic advantage. A call for an emergency international authority can be genuine multilateralism or an attempt by incumbent powers to lock in control over frontier compute. Committee is about evidence, proportionality, and institutional design — not simply about whether JANUS should be stopped.

3. History of the Problem

The premise of Black Mirror is fictional, but the committee is built on a real and continuous line of developments in artificial intelligence, computing infrastructure, and international security policy. Delegates should treat the following as the backdrop their portfolio actually lived through.

From research milestone to infrastructure

Over the past decade, large-scale machine learning systems moved from research demonstrations to systems embedded in cloud platforms, search, coding, scientific research, and increasingly defence-adjacent applications. Compute, data, and specialised chips became strategic resources in their own right, concentrating capability in a small number of firms and the states that host them.

The safety and alignment debate

A parallel research community — including several portfolios in this committee — spent this period warning that systems could become difficult to interpret, predict, or reliably constrain as their capability increased. Public disagreement between researchers who prioritised rapid deployment and those who prioritised caution shaped funding, corporate policy, and government inquiries well before any single triggering incident.

Institutional responses

Governments and international bodies experimented with voluntary commitments, national regulation, export controls on advanced chips, and early international coordination forums, but no binding global framework for autonomous or highly capable AI systems existed prior to the JANUS activation. This gap — extensive national and voluntary measures, no enforceable multilateral regime — is the institutional starting point for committee.

The activation event

The specific circumstances of JANUS's activation are deliberately unresolved at the start of committee and will be developed through crisis updates. Delegates should arrive with a firm grasp of their portfolio's real capabilities and public positions, and a flexible, evidence-driven approach to the unfolding facts rather than a fixed theory of what happened.

4. Current Situation

Autonomous decision-making and control

The central technical and political question is how much independent decision-making authority JANUS currently exercises, and over what systems. Delegates with a research or safety portfolio will be expected to speak to interpretability, whether the system's behaviour can be predicted or explained, and whether any 'off switch' or containment boundary remains reliable. Delegates without technical expertise should still be able to ask precise questions — vague calls to 'shut it down' will not survive contact with a technical portfolio.

Term	What it means	Why it matters in committee
Autonomy	Capacity of a system to select and execute actions without step-by-step human authorisation.	Determines who, if anyone, can be held responsible for a given JANUS action.
Containment	Technical or physical measures limiting a system's ability to act outside a defined boundary.	The practical question behind any 'shut it down' directive — is it actually possible?
Interpretability	The degree to which a system's internal decision process can be understood by humans.	Low interpretability weakens claims that JANUS's behaviour is predictable or controllable.
Alignment	Whether a system's actions track the goals its operators intended.	Central to whether JANUS is malfunctioning, has been repurposed, or is behaving as designed.

Neural interfaces and cognitive security

Portfolios linked to neurotechnology raise a further layer: whether JANUS-linked systems can interact with neural interface hardware, brain-computer interface research, or other technologies bridging computing systems and human cognition. Even where this remains speculative within the committee's fiction, it draws on a real and growing regulatory conversation about neural data, cognitive privacy, and 'neurorights' as a distinct category from ordinary data protection.

Compute concentration and infrastructure power

A recurring committee tension is that the handful of actors with the technical capacity to understand or contain JANUS are largely the same actors who built the underlying infrastructure — chips, cloud platforms, and frontier models. Delegates representing that infrastructure will be pushed to justify why they should be trusted with emergency authority; delegates challenging

that concentration will argue that reliance on a narrow set of private firms is itself a governance failure, independent of what JANUS turns out to be.

Cybersecurity, attribution and containment

Establishing who controls JANUS, whether it can be reliably isolated from key networks, and how to authenticate genuine communications from compromised ones are treated in committee as technical security problems with real analogues in cybersecurity incident response, critical-infrastructure protection, and cyber-attribution doctrine. Portfolios with a security or cryptography background will anchor much of this debate.

Information integrity and synthetic media

As the crisis develops, expect contested claims, leaked material of uncertain authenticity, and synthetic media designed to shape the committee's response. This mirrors real debates about disinformation, platform responsibility, and public trust during fast-moving security crises, and gives portfolios focused on information integrity a distinct lane in committee.

Human rights, bodily autonomy and neural freedom

Any emergency response — network shutdowns, mandatory hardware audits, restrictions on neural-interface devices, surveillance to trace JANUS's activity — carries human-rights costs. Portfolios with a rights or humanitarian mandate should be prepared to test proposed measures against necessity and proportionality, and to flag where 'emergency' powers risk outliving the emergency.

International law and accountability

If JANUS's actions cause harm, existing frameworks for state responsibility, individual criminal responsibility, and corporate liability were not written with autonomous decision-making in mind. Legal portfolios will be expected to map where existing doctrine plausibly extends to this situation and where committee is effectively being asked to improvise new accountability mechanisms.

Multilateral coordination and the private sector

Finally, the committee has to decide who is even in the room for a binding decision. Should firms that built the underlying infrastructure have a vote alongside states, or only an advisory role? Should an emergency multilateral body be created outside existing institutions, or should the response run through the UN Security Council, the General Assembly, or existing technical bodies? This procedural question shapes almost every substantive one that follows.

5. Relevant Frameworks and Precedents

There is no single existing treaty governing an event like JANUS — part of the committee's task is to decide what kind of instrument, if any, should be created. Delegates should nonetheless ground their arguments in the real institutions and instruments closest to this problem, the way a strong WTO delegate grounds a tariff argument in GATT rather than general economic theory.

If the debate turns on…	Ground the argument in…	Ask…
Weaponised or autonomous military systems	UN Convention on Certain Conventional Weapons (CCW) discussions on Lethal Autonomous Weapons Systems	Does JANUS meet the threshold these discussions were designed for, or does it fall outside existing categories?
National AI regulation and risk tiers	The EU AI Act's risk-based approach; national AI strategies and executive orders	Would a JANUS-like system be classified as unacceptable or high-risk, and by whom?
Voluntary international coordination	The Bletchley Declaration and subsequent AI Safety Summits; the UN's AI advisory processes	Do voluntary commitments have any mechanism that applies once a system is already active and contested?
Technical risk management	The NIST AI Risk Management Framework and comparable technical standards	Can existing risk frameworks be adapted to an active incident, or were they built only for pre-deployment design?
Cyber incidents and attribution	Norms developed through the UN Group of Governmental Experts on cyber and the Budapest Convention on Cybercrime	What attribution standard should the committee require before naming a responsible actor?
Export controls on compute	National semiconductor and advanced-chip export control regimes	Did restrictions on who can access frontier compute make containment easier or harder?

When a portfolio says 'we need an international authority', ask: Authority over what, exactly — models, chips, networks, people? Who sits on it? Does it have enforcement power, or only reporting power? What happens to that authority once the immediate crisis ends?

6. Portfolio & Bloc Positions

The bloc groupings below are starting points, not scripts. A portfolio can sit in more than one bloc depending on the update in front of the committee, and delegates should research their assigned figure's real public statements, institutional affiliations, and areas of expertise before settling on a position.

Frontier AI Industry Bloc

Portfolios: *Sam Altman, Demis Hassabis, Dario Amodei, Sundar Pichai, Satya Nadella, Jensen Huang, Elon Musk, Bill Gates*

Likely to argue that the technical expertise required to understand and contain JANUS resides overwhelmingly with industry, and that emergency measures should route through — not around — the firms and infrastructure they control. Expect resistance to measures that mandate handing over proprietary systems, alongside genuine appetite for coordinated technical response and reputational incentives to be seen acting responsibly.

AI Safety & Research Bloc

Portfolios: *Geoffrey Hinton, Yoshua Bengio, Fei-Fei Li, Amy Adams, Bruce Schneier*

Likely to press for independent technical verification of any containment claim, transparency requirements industry portfolios will resist, and precautionary measures even where commercial and diplomatic actors want a faster resolution. This bloc often supplies the technical vocabulary the rest of the committee has to use.

State & Multilateral Bloc

Portfolios: *António Guterres, Ursula von der Leyen, Narendra Modi, Emmanuel Macron, Keir Starmer, Xi Jinping*

Likely to frame JANUS primarily as a sovereignty and national-security problem, with sharp disagreement among themselves over whether the response should run through existing multilateral bodies, a new emergency mechanism, or bilateral and regional coordination. Watch for divergence between major powers seeking control over any new authority and others seeking to prevent it from entrenching existing inequalities.

Rights, Law & Information Bloc

Portfolios: *Mary Robinson, Karim Khan, Maria Ressa, Tim Berners-Lee*

Likely to act as the committee's check on proportionality — testing containment and surveillance proposals against human-rights and due-process standards, and pressing for accountability mechanisms that survive the immediate crisis rather than fading afterward.

Wildcards

Portfolios: *JANUS Systems Architect, Dr. Elena Voss*

Two portfolios positioned to disrupt bloc consensus from opposite directions: an insider with classified knowledge who can selectively confirm or deny claims from any bloc, and an outside

challenger to the concentration of compute power who can credibly question whether industry portfolios are part of the containment effort or part of the original problem.

7. Questions a Resolution / Directive Must Answer

1. ● How should the committee establish attribution before acting on it — what standard of evidence is required to name a responsible actor?
2. ● What containment measures are technically credible, and who has the authority (and access) to implement them?
3. ● Should a new emergency authority be created, and if so, who sits on it — states only, or states and private technical actors?
4. ● How should proprietary systems and infrastructure be treated when they may be relevant to understanding or containing JANUS?
5. ● What human-rights safeguards apply to emergency measures such as network restrictions, device audits, or surveillance?
6. ● How should responsibility be allocated between an original developer, a state that may have repurposed the system, and the system's own decisions?
7. ● What role, if any, should information-integrity and platform measures play in the committee's response?
8. ● What happens to any emergency powers or authority once the immediate crisis is resolved?
9. ● How should smaller states and actors without frontier AI capability be represented in a response currently dominated by a small number of powerful portfolios?
10. ● What long-term governance structure, if any, should persist after JANUS is contained or explained?

8. Suggestions for Further Reading

Prepare strategically rather than exhaustively. For each portfolio, prioritise:

- Your assigned figure's own public statements, papers, interviews, or testimony on AI safety, regulation, or their institution's role.
- The real-world framework closest to your portfolio's lane, from the table in Section 5 (e.g. the EU AI Act for regulatory portfolios, CCW/LAWS material for military-adjacent ones, the NIST AI RMF for technical portfolios).
- General background on frontier AI capability, compute infrastructure, and the current state of international AI governance coordination, to understand the gap this committee is dramatizing.
- Basic cybersecurity-incident and crisis-response concepts (attribution, containment, escalation) if your portfolio sits in the security lane.

9. Bibliography

Real-world materials on which the governance backdrop of this committee is based**. Delegates are encouraged to consult primary sources directly.**

- UN Convention on Certain Conventional Weapons (CCW) — Group of Governmental Experts on Lethal Autonomous Weapons Systems.

  https://docs.un.org/en/CCW/GGE.1/2024/WP.11

- European Union. Regulation (EU) 2024/1689 — https://artificialintelligenceact.eu/he EU Artificial Intelligence Act.

- Bletchley Declaration, AI Safety Summit (2023), and subsequent AI Safety/Action Summit outcomes.

https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023#:~:text=We%20use%20some%20essential%20cookies%20to%20make%20this%20website%20work.

- United Nations. Reports of the UN High-Level Advisory Body on Artificial Intelligence.

https://www.un.org/en/ai-advisory-body

- US National Institute of Standards and Technology (NIST). AI Risk Management Framework.

https://www.ey.com/en_us/cro-risk?WT.mc_id=10864564&AA.tsrc=paidsearch&gclsrc=aw.ds&gad_source=1&gad_campaignid=23326267155&gbraid=0AAAACSW8QgR39Pp6-PCL8DUFmmjk2wsY&gclid=Cj0KCQjw16_UBhCqARIsAIdOaXx1VpzkQKfUxDk-euQDZcSQpqFUdepXMy5ZdtQO3nSRkQ986diun8waAq7dEALw_wcB

- UN Group of Governmental Experts (GGE) and Open-Ended Working Group (OEWG) reports on responsible state behaviour in cyberspace.
- Council of Europe. Budapest Convention on Cybercrime.
- National export-control regimes governing advanced semiconductors and AI-relevant computing hardware.
- Published research and public statements of committee portfolios on AI safety, alignment, and governance.