

2 JUILLET 2026

Droit international et intelligence artificielle

Une approche infrastructurelle

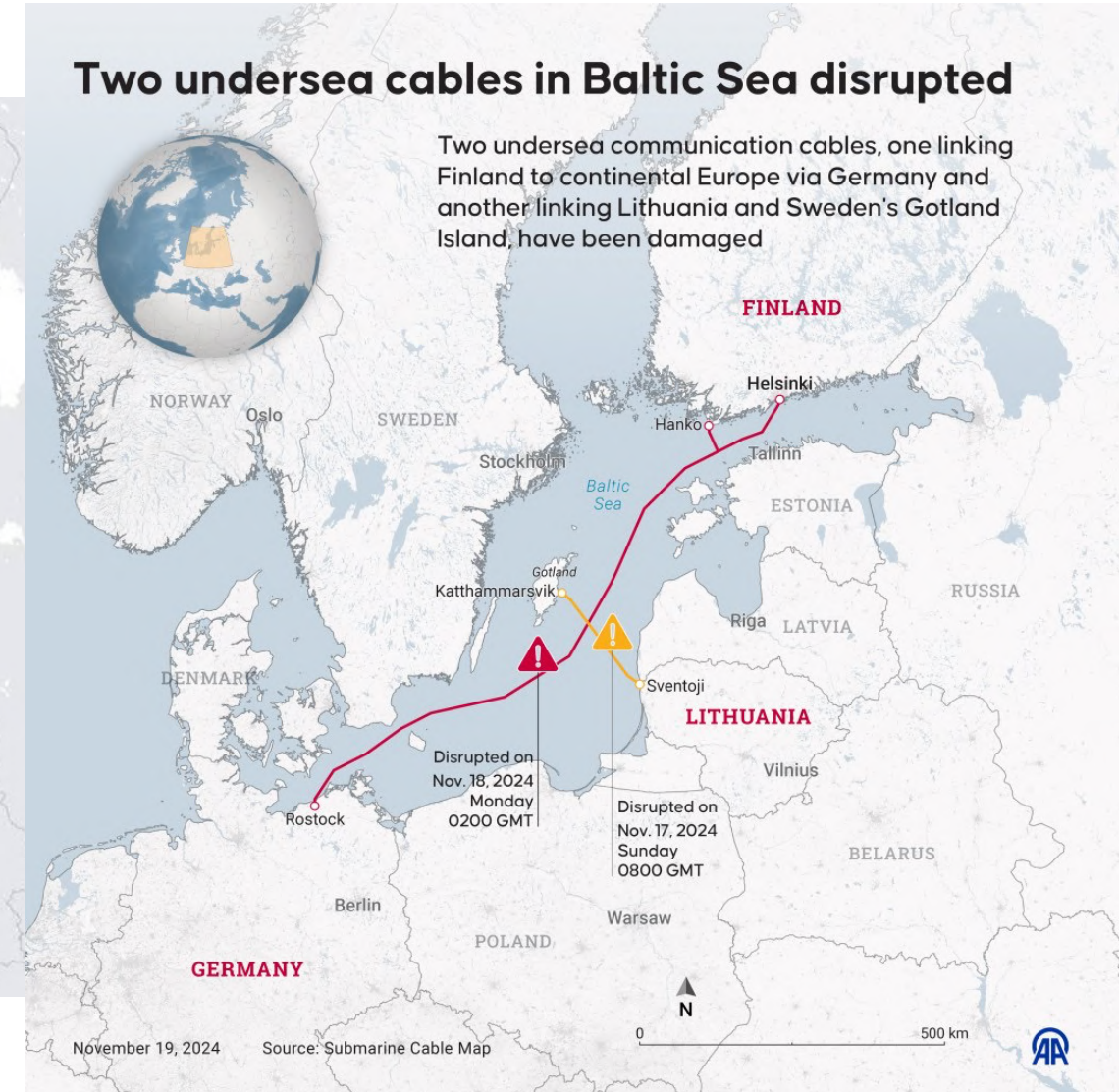
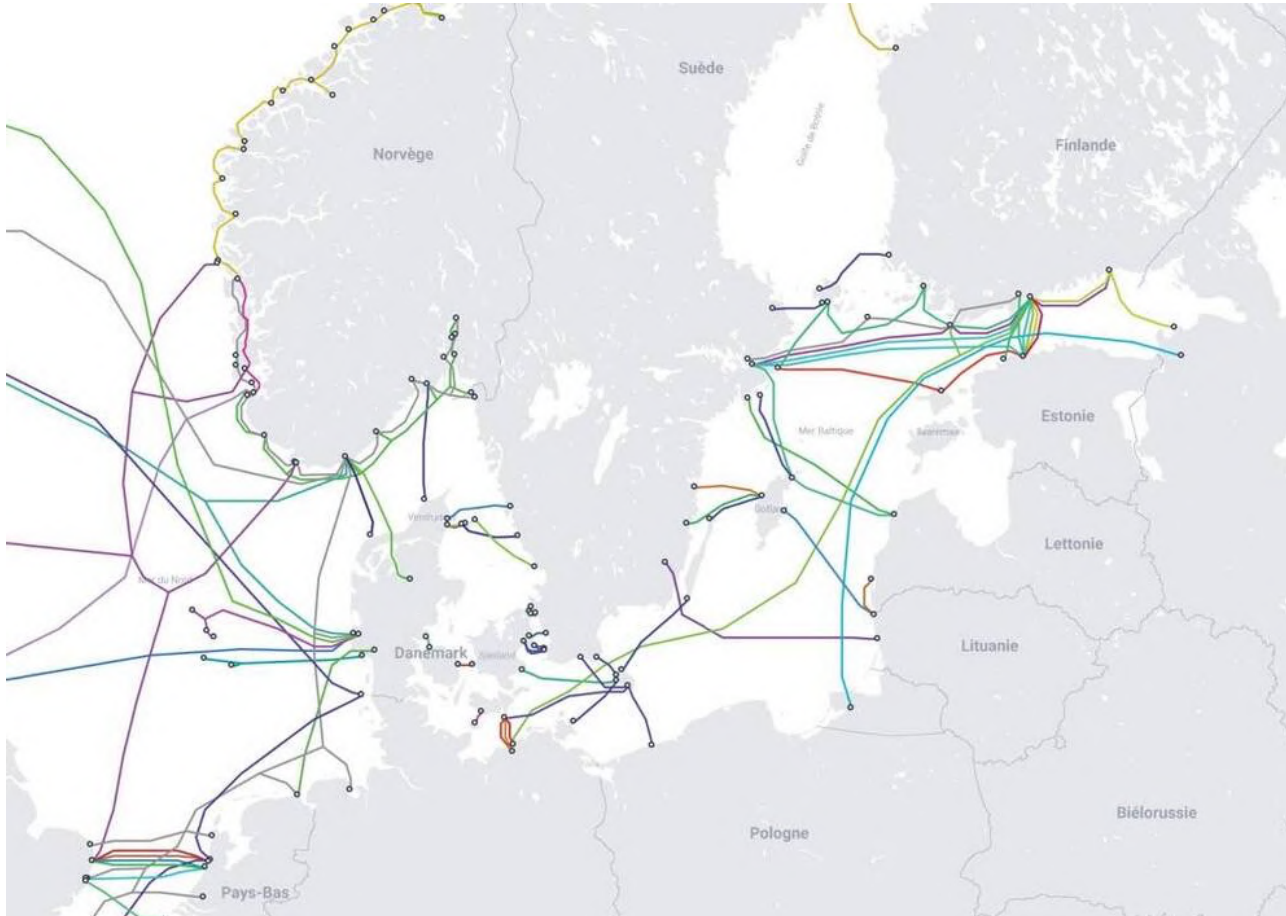
Raphaël Maurel — MCF HDR de droit public, Université Bourgogne Europe (CREDIMI), Institut universitaire de France

Mises en situation

La guerre à Gaza



Les câbles sous-marins en mer Baltique



L'implantation des *datacenters*

Highlander – China's Underwater Data Center Cluster (Hainan)

- **Location:** South China Sea, near Hainan Island
- **First capsule:** Deployed March 2023; **Second:** Deployed February 2025
- ~1,200 servers as of 2025
- Powered by mainland grid; offshore wind planned
- Operational, expanding to 100 capsules within 5 years



La fin de (la) Fable



Announcements

Statement on the US government directive to suspend access to Fable 5 and Mythos 5

12 juin 2026

The US government, citing national security authorities, has issued an export control directive to suspend all access to Fable 5 and Mythos 5 by any foreign national, whether inside or outside the United States, including foreign national Anthropic employees. The net effect of this order is that we must abruptly disable Fable 5 and Mythos 5 for all our customers to ensure compliance. **Access to all other Anthropic models will not be affected.**

We received the directive from the government today at 5:21pm (ET). The letter did not provide specific details of its national security concern. Our understanding is that the government believes it has become aware of a method of bypassing, or “jailbreaking” Fable 5. We reviewed a demonstration of this specific technique being used to identify a small number of previously known, minor vulnerabilities. These vulnerabilities all appear relatively simple, and we have found that other publicly-available models are able to discover them as well without requiring a bypass.

Ce que ces situations ont en commun



Une matérialité

Câbles, data centers immergés, installations en mer : l'IA a un corps physique localisable.



Une territorialité

Chaque maillon se situe dans un espace (territoire, mer, orbite) doté d'un régime juridique.



Une vulnérabilité

L'infrastructure peut être identifiée et atteinte, donc protégée ou instrumentalisée.



Des questions de droit

...auxquelles il est apporté peu de réponses.

***Thèse : on ne peut analyser ni l'IA, ni le droit de l'IA en partant des seules « applications ».
Il faut partir de l'infrastructure qui les rend possibles.***

CHAPITRE INTRODUCTIF

Qu'est-ce que l'intelligence artificielle ?

Un objet ancien, une définition récente

1956

Le terme « artificial intelligence » est forgé lors d'un colloque au Dartmouth College (proposition de J. McCarthy, M. Minsky, N. Rochester, C. Shannon).

1950–1980

Approches symboliques (logique, systèmes experts) : l'humain encode connaissances et règles.

2010→

Apprentissage automatique et modèles statistiques : la « performance » émerge de l'entraînement sur des données.

2019→

Le législateur national, européen, international s'empare de l'objet économique. L'OCDE définit le « système d'IA » plutôt que l'« IA ».

On retiendra surtout que, faute de consensus sur « l'IA », les instruments juridiques définissent un objet plus opératoire : le système d'IA.

Définition de l'OCDE (2019, révisée en 2023)

“ *Un système d'IA est un système fondé sur une machine qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des résultats tels que des prédictions, des contenus, des recommandations ou des décisions qui peuvent influencer des environnements physiques ou virtuels.*

OCDE, Recommandation du Conseil sur l'IA (OECD/LEGAL/0449), définition du « système d'IA » révisée le 8 nov. 2023

Du *soft* au *hard law* (regional) – l'UE



« système d'IA », un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer des environnements physiques ou virtuels.

Règlement (UE) 2024/1689 (RIA), art. 3, point 1) — JO L 2024/1689, 12 juill. 2024

OCDE	Recommandation OECD/LEGAL/0449, déf. révisée le 8 nov. 2023 — soft law fondatrice.
Union européenne	Règlement (UE) 2024/1689, art. 3(1) — hard law, applicable par paliers jusqu'au 2 août 2026.
Conseil de l'Europe	Convention-cadre (STCE n° 225), art. 2 — premier traité contraignant, ouvert à la signature le 5 sept. 2024 (Vilnius).

Du *soft* au *hard law* (regional) – Le Conseil de l'Europe

“ un système automatisé qui, pour des objectifs explicites ou implicites, déduit, à partir d'entrées reçues, comment générer des résultats en sortie tels que des prévisions, des contenus, des recommandations ou des décisions qui peuvent influencer sur des environnements physiques ou virtuels. Différents systèmes d'intelligence artificielle présentent des degrés variables d'autonomie et d'adaptabilité après déploiement.

Convention-cadre du Conseil de l'Europe sur l'IA et les droits de l'homme, la démocratie et l'État de droit, article 2

OCDE	Recommandation OECD/LEGAL/0449, déf. révisée le 8 nov. 2023 — soft law fondatrice.
Union européenne	Règlement (UE) 2024/1689, art. 3(1) — hard law, applicable par paliers jusqu'au 2 août 2026.
Conseil de l'Europe	Convention-cadre (STCE n° 225), art. 2 — premier traité contraignant, ouvert à la signature le 5 sept. 2024 (Vilnius).

Définition des “technologies d’IA”

“ *toute technologie ou tout savoir-faire spécifiquement lié à un système fondé sur une machine qui est conçu pour fonctionner avec des niveaux d’autonomie variables, qui peut faire preuve d’adaptabilité après le déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des données qu’il reçoit, comment générer des résultats tels que des prédictions, des contenus, des recommandations ou des décisions susceptibles d’influencer les environnements physiques ou virtuels («système d’IA»), utilisées pour:*

- a) les modèles d’IA à usage général tels qu’ils sont définis à l’article 3, point 63), du règlement (UE) 2024/1689 du Parlement européen et du Conseil ou les systèmes d’IA fondés sur de tels modèles adaptés au développement d’applications spatiales ou de défense; ou*
- b) les modèles d’IA à usage général présentant un risque systémique au sens de l’article 51 du règlement (UE) 2024/1689 ou les systèmes d’IA fondés sur de tels modèles.*

Règlement (UE) 2026/1386 du Parlement européen et du Conseil du 17 juin 2026 relatif au filtrage des investissements étrangers dans l’Union et abrogeant le règlement (UE) 2019/452

Système d'IA ≠ « technologies d'IA »



Le système d'IA

- Objet défini par l'OCDE, le RIA et la Convention-cadre du CdE.
- Approche très fonctionnelle : ce que le système fait (inférer, générer des sorties).
- Saisi par le droit de la concurrence (régulation des marchés) et dans une moindre mesure par les droits fondamentaux.



Les technologies d'IA

- Notion presque absente du RIA, mais présente ailleurs : semi-conducteurs, calcul, modèles.
- Approche par les intrants : matériaux, puces, infrastructures critiques.
- Saisie par le contrôle des investissements et des exportations (sécurité économique, défense).

Qu'est-ce qu'une approche infrastructurelle ?

- Trois approches dominantes de regulation en droit positif et dans la doctrine : par les risques, par les facilités essentielles, par les droits fondamentaux.
- Une proposition : repartir de la base matérielle qui détermine tout le reste.
- De l'extraction minière aux satellites : quelle chaîne de l'IA ?

L'approche par les risques (dominante)

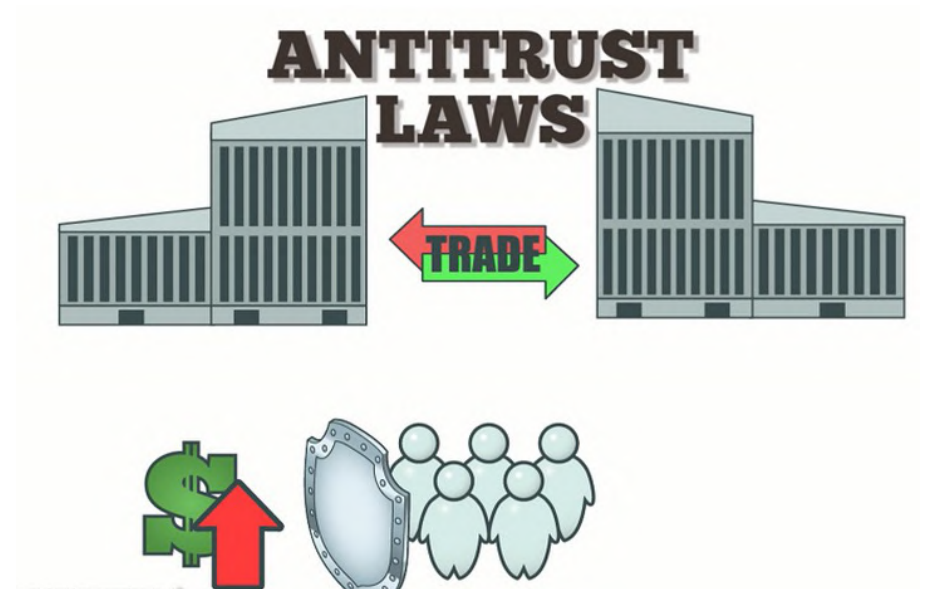
- Logique du Règlement (UE) 2024/1689 : graduer les obligations selon le niveau de risque du système + introduction au dernier moment de la notion de “modèle à usage général” pour tenir compte de l’irruption de ChatGPT.
- Quatre niveaux : risque inacceptable (pratiques interdites), élevé, limité (transparence), minimal.
- Centrée sur la mise sur le marché et l'usage ; raisonne par application et par finalité.
- Force opérationnelle, protège la capacité d’innovation, favorise l’acceptabilité sociale par la mention des droits fondamentaux et de l’éthique.
- Limite : suppose l'infrastructure déjà là, comme une donnée.

A. Latil, *Le droit du numérique. Une approche par les risques*, Dalloz, 2ème éd., 2024.



L'approche par les facilités essentielles

- Doctrine des essential facilities, héritée du droit *antitrust* états-unicien.
- Idée : certaines infrastructures sont indispensables et conditionnent l'accès au marché.
- Transposée à l'IA : calcul, modèles de fondation, jeux de données sont des « facilités essentielles ».
- Enjeu : garantir un accès équitable et non discriminatoire aux intrants critiques de l'IA.
- Limites : réception très limitée car le droit de la concurrence s'arrête là où les enjeux de défense commencent.



L'approche par les droits fondamentaux

- Place les droits de la personne au centre : dignité, non-discrimination, vie privée, recours...
- Logique portée par la Convention-cadre du Conseil de l'Europe (STCE n° 225, Vilnius, 2024).
- Complémentaire de l'approche par les risques, dont elle fixe la finalité et les limites.
- Limites : se borne généralement à une approche de type principiste / principielle mal traduite en droit positif.



L'approche infrastructurelle / par les infrastructures : partir de la base

On ne régle pas l'IA en partant des seules applications. On part de l'infrastructure physique qui détermine tout le reste.



Chaque maillon est saisi par des régimes juridiques distincts ; c'est leur articulation qui constitue le « droit international de l'IA ».

Plan

Chapitre introductif
Chapitre 1. Extraire
Chapitre 2. Stocker
Chapitre 3. Connecter
Chapitre 4. Utiliser

CHAPITRE 1

Extraire – matériaux de l'IA

- De quels matériaux a-t-on besoin pour calculer et donc développer l'IA ?
- Quel droit international des ressources et des chaînes d'approvisionnement ?

La matérialité minérale du calcul

- Puces et serveurs reposent sur des métaux critiques et des terres rares.
- La production est géographiquement concentrée = dépendances et vulnérabilités stratégiques.
- La « dématérialisation » de l'IA repose en réalité sur une intensification extractive.



Le cadre international de l'extraction des ressources

- Souveraineté permanente sur les ressources naturelles
- Régimes spécifiques : grands fonds marins (Autorité internationale des fonds marins, CNUDM partie XI)
- Encadrement des chaînes d'approvisionnement et devoirs de diligence (minéraux de conflits)
- Droit des investissements ?

Exploration for minerals in the Area



Sécurité et matériaux

- Les matériaux critiques sont saisis comme enjeux de sécurité économique mais aussi de sécurité internationale tout court
- Contrôles d'investissements et d'exportations sur les intrants stratégiques
- L'extraction est un objet de politique étrangère, de droit international économique mais aussi de droit international de la sécurité collective



[\[French version\]](#)

The Government of the United States of America ("United States") and the Government of the Democratic Republic of the Congo ("DRC"), each a "Party" and hereinafter collectively referred to as the "Parties,"

Acknowledging the longstanding partnership between the United States and the DRC, and the mutual commitment to fostering peace, stability, and long-term economic development in the region;

Recognizing the Parties' shared interest in furthering cooperation to advance economic and resource security, and to promote long-term development for the benefit of both nations;

Puces et double usage

- Biens à double usage (Wassenaar, Règlement 2021/821)
- *Chips Acts*
- Qualification comme enjeu de sécurité nationale et internationale



41 Wassenaar Arrangement
Participating States



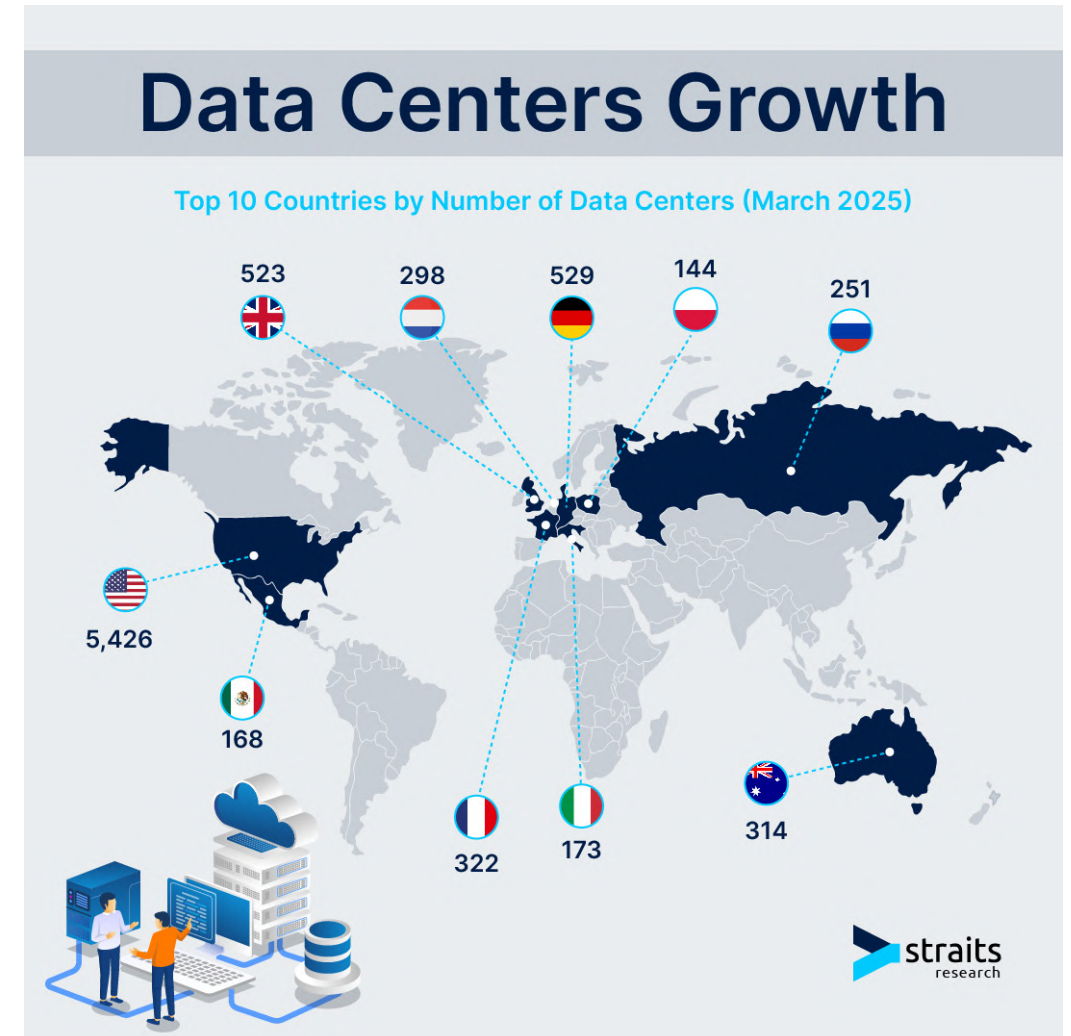
CHAPITRE 2

Stocker – données et centres de données

- Comment stocke-t-on les données et quel droit international s'applique ?
- Centres de données terrestres, immergés, orbitaux ?

Le centre de données comme territoire virtuel ?

- Le stockage se réalise dans des centres de données localisés : le “cloud” n’existe pas.
- La localisation détermine la loi applicable et la souveraineté en jeu.
- Tendances à la « localisation des données » (data localization) et leurs limites.



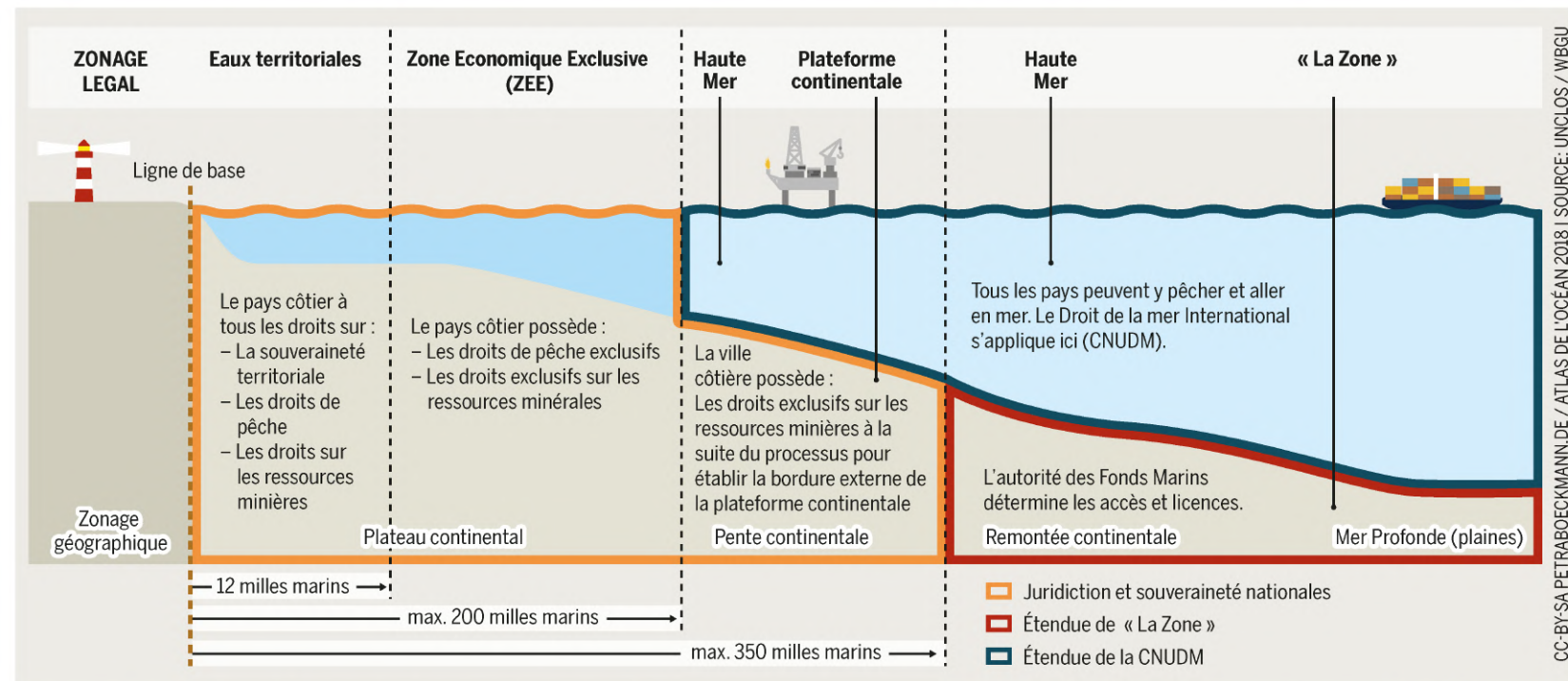
Des centres de données en mer ?

“ Dans la zone économique exclusive, l'État côtier a [...] juridiction [...] en ce qui concerne [...] la mise en place et l'utilisation d'îles artificielles, d'installations et d'ouvrages [...] ”

CNUDM (Montego Bay, 10 déc. 1982), art. 56, §1, b), i)

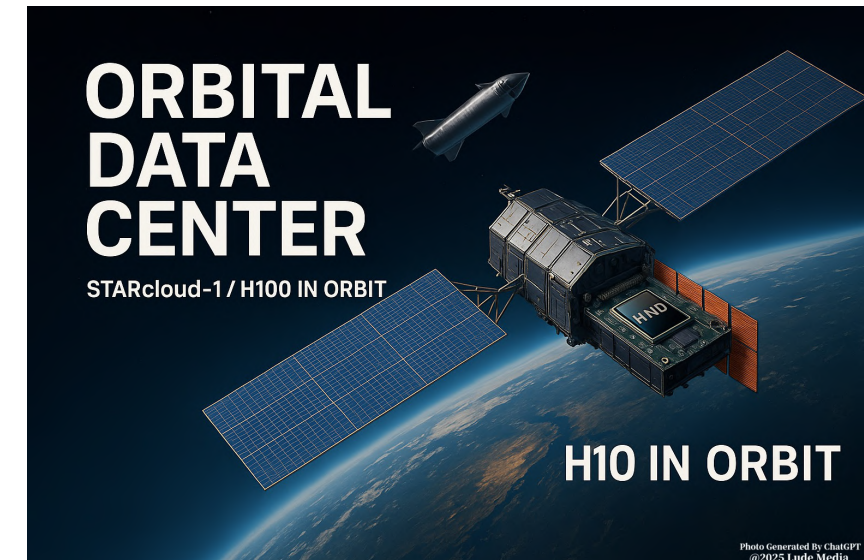
- Centres immergés : qualification comme installation, ouvrage ou “autre utilisation” de la mer ?
- Régime variable selon la zone : mer territoriale, ZEE, plateau continental, haute mer

Comment pensent les juristes — Zones maritimes et droit international de la mer



Des centres de données en orbite ?

- Un engin spatial au sens du Traité de 1967
- Application du droit de l'espace extra-atmosphérique



Transferts de données et souveraineté

- Transférer une donnée d'un endroit de stockage à un autre = franchir des frontières.
- Encadrement sectoriel des transferts internationaux de données à caractère personnel

Réf. — RGPD (règlement (UE) 2016/679), chap. V ; jurisprudence sur les transferts internationaux.



Ressources et empreinte du stockage

- Les centres de données consomment massivement énergie + eau de refroidissement.
- D'où l'immersion : utiliser la mer comme dissipateur thermique (= “soupe de poissons” ?)
- Articulation avec le droit de l'environnement et du climat ?

COUR INTERNATIONALE DE JUSTICE

ANNÉE 2025

2025
23 juillet
Rôle général
n° 187

23 juillet 2025

OBLIGATIONS DES ÉTATS EN MATIÈRE DE CHANGEMENT CLIMATIQUE

Compétence et pouvoir discrétionnaire.

Questions portées devant la Cour revêtant un caractère juridique — Cour ayant compétence pour donner l'avis consultatif demandé — Absence de raisons décisives qui justifieraient que la Cour refuse de donner l'avis consultatif demandé.

* *

Contexte dans lequel la résolution 77/276 a été adoptée — Conséquences graves et profondes des changements climatiques — Mesures les plus pertinentes prises à l'échelon international entre 1968 et 2015 pour cerner les risques de manière à protéger le système climatique et d'autres composantes de l'environnement — Contexte scientifique pertinent — Rapports du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) constituant les meilleures données scientifiques disponibles sur les causes, la nature et les conséquences des changements climatiques — Conclusions du GIEC sur les causes et les conséquences des changements climatiques — Émissions de gaz à effet de serre (GES) et changements climatiques — Mesures d'atténuation et d'adaptation.

* *

CHAPITRE 3

Connecter – câbles et satellites

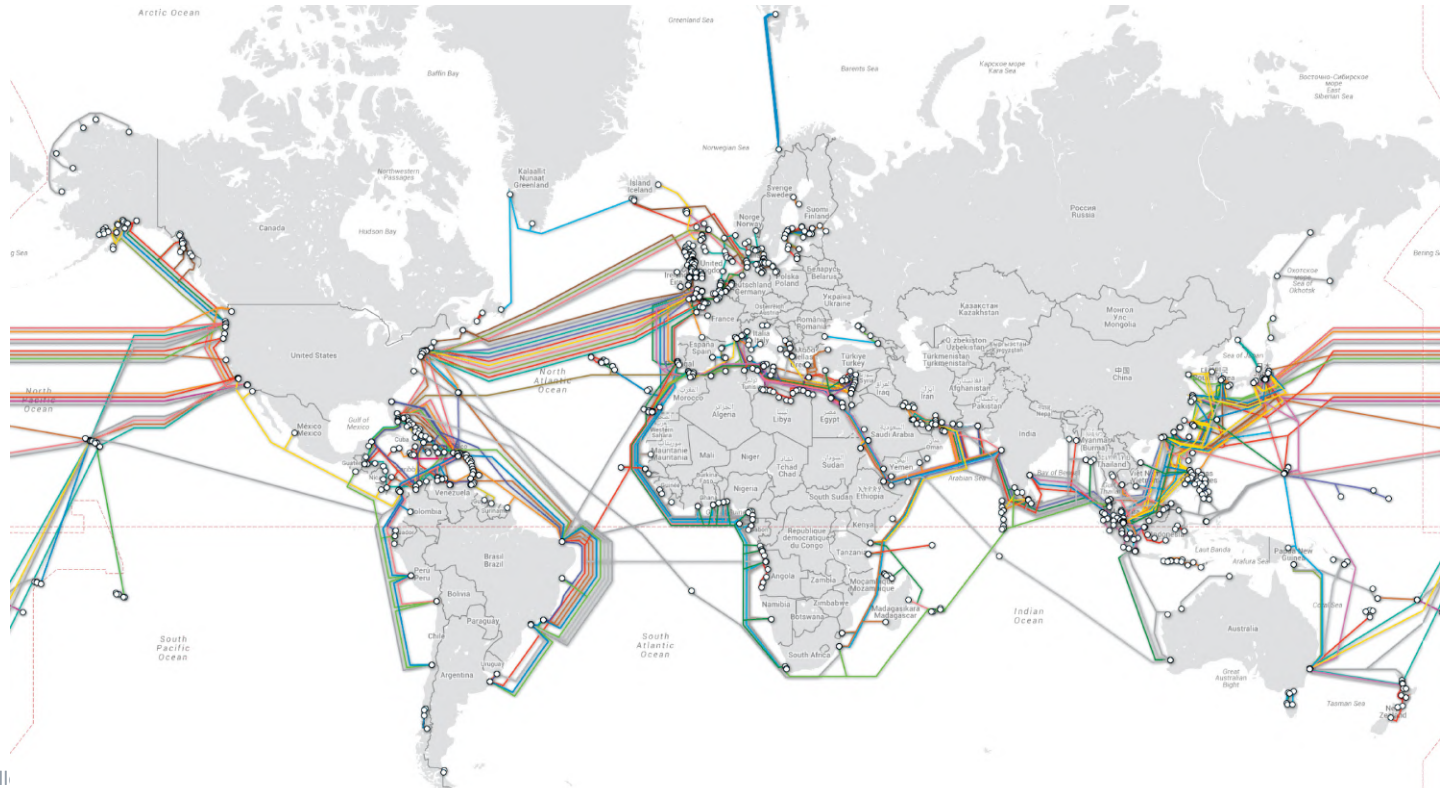
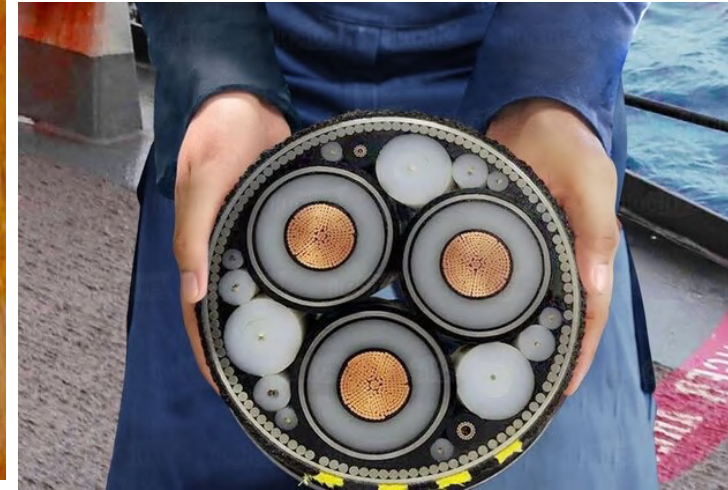
- Comment circule la donnée et la puissance de calcul ?
- Câbles sous-marins, câbles terrestres, satellites : emprunt au droit des infrastructures numériques
- Le cyberspace a une géographie physique !

Les câbles sous-marins

“ Tous les États ont le droit de poser des câbles et des pipelines sous-marins sur le plateau continental.

CNUDM (1982), art. 79, §1

- L'essentiel du trafic international transite par des câbles sous-marins
- Convention pour la protection des câbles sous-marins (Paris), 14 mars 1884, 36 États parties
- Convention de Genève sur la haute mer, 29 avril 1958
- Régime de la CNUDM : liberté de pose, mais protection lacunaire contre les atteintes.



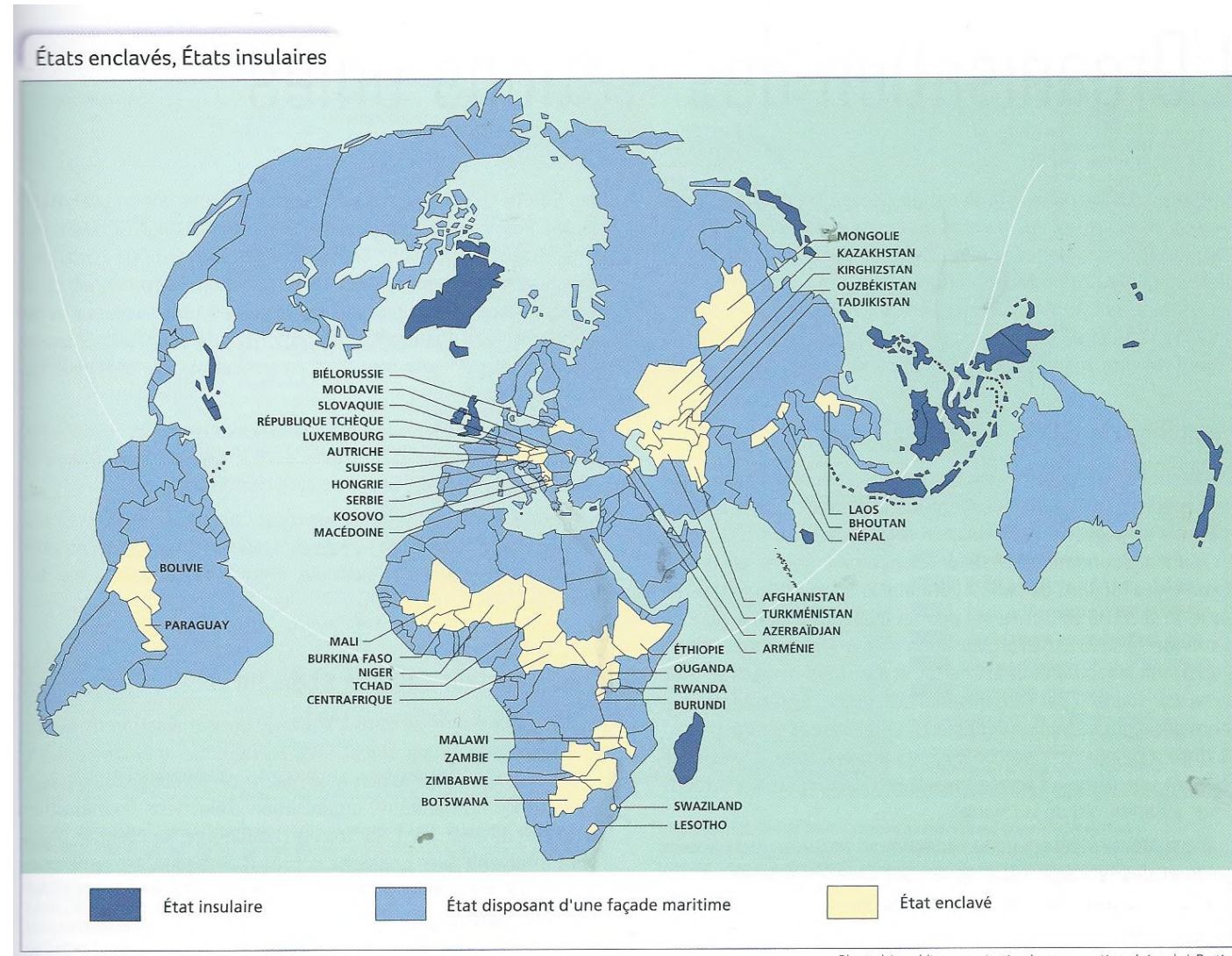
Protéger les câbles : un régime lacunaire

- Obligation d'incriminer la rupture volontaire (art. 113 CNUDM), mais effectivité limitée.
- Difficultés d'imputation et de compétence au-delà de la mer territoriale.
- Débats actuels



Les câbles terrestres

- La donnée circule aussi par des dorsales terrestres et des points d'atterrissement / atterrissage
- Souveraineté et contrôle des infrastructures de communication
- Accords bilatéraux (ex : accord privé Mali-Guinée juin 2024) et régionaux
- Risques de cybersécurité



Les satellites

“ *L'espace extra-atmosphérique [...] ne peut faire l'objet d'appropriation nationale par proclamation de souveraineté, ni par voie d'utilisation ou d'occupation, ni par aucun autre moyen.*

Traité sur les principes régissant les activités des États en matière d'exploration et d'utilisation de l'espace extra-atmosphérique, y compris la Lune et les autres corps célestes, 27 janv. 1967, art. II

- Constellations en orbite basse = nouvelle couche de connectivité et de calcul.
- Traité de 1967 et droit de l'espace (liberté d'exploration, non-appropriation, responsabilité)
- L'orbite, une ressource limitée (encombrement et difficultés de coordination des fréquences)

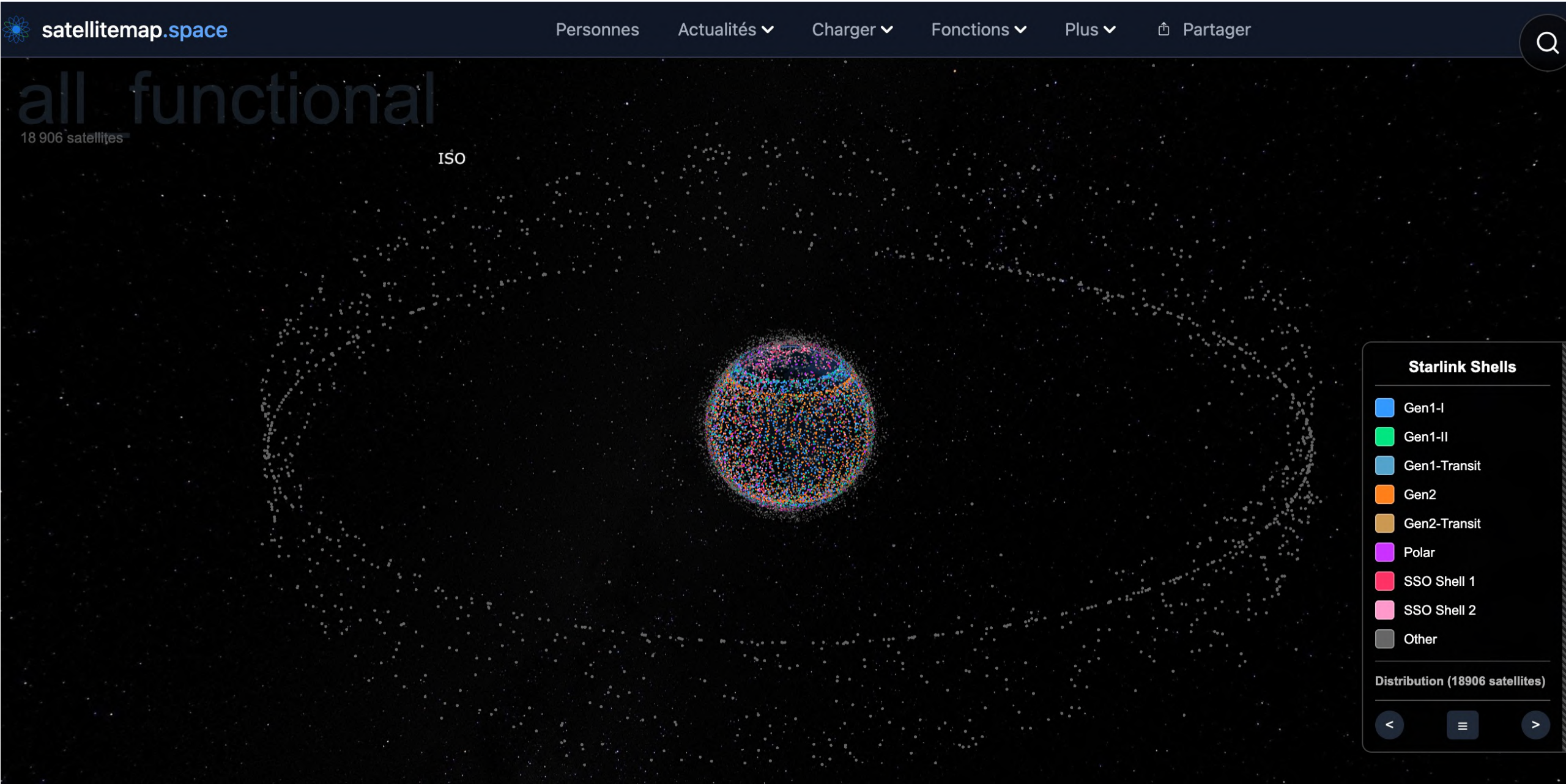
2222 (XXI). Traité sur les principes régissant les activités des États en matière d'exploration et d'utilisation de l'espace extra-atmosphérique, y compris la Lune et les autres corps célestes

L'Assemblée générale,

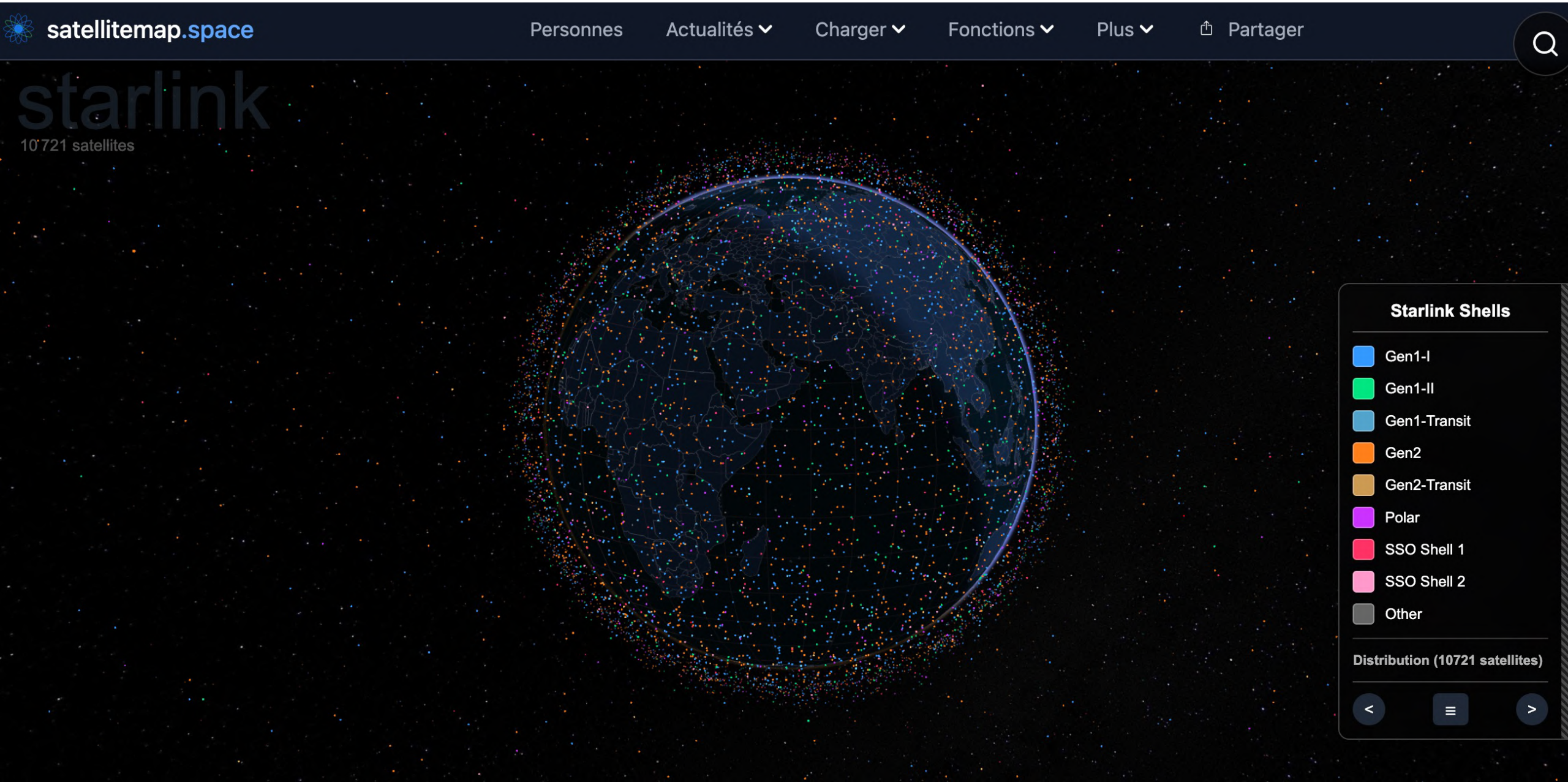
Ayant examiné le rapport du Comité des utilisations pacifiques de l'espace extra-atmosphérique sur ses travaux de l'année 1966¹⁰, et en particulier l'œuvre accomplie par le Sous-Comité juridique à sa cinquième session, tenue à Genève du 12 juillet au 4 août et à New York du 12 au 16 septembre,

Notant en outre les progrès accomplis grâce à des consultations ultérieures entre les États Membres de l'Organisation des Nations Unies,

Total satellites – 1er Juillet 2026 : 18906



Total Starlink – 1er Juillet 2026 : 10721



Débris 1er Juillet 2026 : 12552



debris

12 552 satellites

KM-A motor debris
1977-012

Starlink Shells

-  Gen1-I
-  Gen1-II
-  Gen1-Transit
-  Gen2
-  Gen2-Transit
-  Polar
-  SSO Shell 1
-  SSO Shell 2
-  Other

Distribution (12552 satellites)



L'interconnexion logicielle

- Au-delà des infrastructures de connexion, les machines reliées doivent être compatibles entre elles pour recevoir et transmettre des données
- Rôle de la standardisation internationale privée / publique
- Enjeux en matière d'intégration des droits de l'homme



CHAPITRE 4

Utiliser – activités mobilisant l'IA

- Le droit international de l'IA emprunte donc largement au droit des infrastructures numériques
- En marge des infrastructures : les activités utilisant l'IA sont encadrées et leur croissance favorisée par le droit international
- Quels textes de droit international ? Quelles activités concernées ?

Conseil de l'Europe – la Convention de Vilnius



L'objet de la présente Convention est de veiller à ce que les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle soient pleinement compatibles avec les droits de l'homme, la démocratie et l'État de droit.

Convention-cadre du Conseil de l'Europe sur l'IA (STCE n° 225), art. 1, §1 — Vilnius, 5 sept. 2024

- Premier traité international juridiquement contraignant portant spécifiquement sur l'IA.
- Technologiquement “neutre” ; centrée sur le cycle de vie des systèmes d'IA.
- Approche fondée sur les droits -> approche de type principielle
- Signée notamment par l'UE, le Royaume-Uni, les États-Unis et Israël.

Chapitre II – Obligations générales

Article 4 – Protection des droits de l'homme

Chaque Partie adopte ou maintient des mesures pour veiller à ce que les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle soient cohérentes avec les obligations de protection des droits de l'homme, telles qu'elles sont définies par le droit international applicable et par son droit interne.

Article 5 – Intégrité des processus démocratiques et respect de l'État de droit

- 1 Chaque Partie adopte ou maintient des mesures visant à garantir que les systèmes d'intelligence artificielle ne sont pas utilisés pour porter atteinte à l'intégrité, à l'indépendance et à l'efficacité des institutions et processus démocratiques, y compris au principe de la séparation des pouvoirs, au respect de l'indépendance de la justice et à l'accès à la justice.
- 2 Chaque Partie adopte ou maintient des mesures qui visent à protéger ses processus démocratiques au cours des activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle, y compris l'accès équitable et la participation des personnes au débat public, ainsi que leur capacité à se forger librement une opinion.

Chapitre III – Principes relatifs aux activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle

Article 6 – Approche générale

Le présent chapitre énonce les principes généraux communs que chaque Partie met en œuvre à l'égard des systèmes d'intelligence artificielle, de manière adaptée à son ordre juridique interne et aux autres obligations nées de la présente Convention.

Article 7 – Dignité humaine et autonomie personnelle

Chaque Partie adopte ou maintient des mesures en faveur du respect de la dignité humaine et de l'autonomie personnelle en ce qui concerne les activités menées dans le cadre du cycle de vie des systèmes d'intelligence artificielle.

Nations Unies – Résolution 78/265 AGNU du 21 mars 2024



Saisir les possibilités offertes par des systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance pour le développement durable.

AGNU, rés. 78/265, adoptée sans vote le 21 mars 2024 (intitulé de la résolution)

- Première résolution de l'Assemblée générale consacrée à l'intelligence artificielle.
- Approche technosolutionniste centrée sur le développement durable (Agenda 2030) et la réduction des fractures numériques.
- Texte non contraignant, adopté par consensus ; champ non militaire.
- Prolongée par la rés. 78/311 (2024) sur le renforcement des capacités, portée par la Chine.

3. *Encourage* les États Membres et invite toutes les parties prenantes de toutes les régions et de tous les pays, chacune dans les limites de sa vocation et de ses responsabilités, notamment le secteur privé, les organisations internationales et régionales, la société civile, les médias, le monde universitaire et les instituts de recherche, ainsi que les milieux techniques et les techniciens, à apporter leur concours à l'élaboration de principes et de cadres de réglementation et de gouvernance permettant de régir des systèmes d'intelligence artificielle qui soient sûrs, sécurisés et dignes de confiance, qui créent un écosystème favorable à tous les niveaux, y compris pour l'innovation, la création d'entreprises et la diffusion des connaissances et des techniques selon des modalités arrêtées d'un commun accord, consciente qu'un partenariat et une coopération efficaces entre les pouvoirs publics et les différentes parties prenantes sont nécessaires pour élaborer ces principes et ces cadres ;

4. *Engage* les États Membres et invite les autres parties prenantes à coopérer avec les pays en développement et à leur apporter une assistance pour favoriser un accès inclusif et équitable aux avantages que procurent la transformation numérique et les systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance,

5. *Souligne* qu'il faut respecter, protéger et promouvoir les droits humains et les libertés fondamentales tout au long du cycle de vie des systèmes d'intelligence artificielle, demande à tous les États Membres et, le cas échéant, aux autres parties prenantes de s'abstenir ou de cesser de se servir des systèmes d'intelligence artificielle qu'il est impossible d'utiliser dans le respect du droit international des droits humains ou qui présentent des risques excessifs pour l'exercice des droits humains, en particulier pour les personnes en situation de vulnérabilité, et réaffirme que les droits dont les personnes jouissent hors ligne doivent également être protégés en ligne, y compris tout au long du cycle de vie des systèmes d'intelligence artificielle ;

Nations Unies – Résolution 78/265 AGNU du 21 mars 2024



Intensifier la coopération internationale en matière de renforcement des capacités dans le domaine de l'intelligence artificielle

AGNU, rés. 78/311, adoptée le 1er Juillet 2024

- Confirmation de l'approche technosolutionniste de type libérale "croissante" : développer l'IA est prioritaire, tout en considérant les risques et le droit international des droits humains

3. *Invite* la communauté internationale à favoriser un environnement propice à la coopération internationale en matière de renforcement des capacités dans le domaine de l'intelligence artificielle et souligne que le droit international doit être respecté, en particulier la Charte des Nations Unies, ainsi que les droits humains et les libertés fondamentales ;

6. *Invite* la communauté internationale à offrir et à favoriser un environnement commercial équitable, ouvert, inclusif et non discriminatoire tout au long du cycle de vie de systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance ;

7. *Encourage* les États Membres à peser les avantages et les risques lorsqu'ils généralisent l'utilisation responsable de solutions numériques, telles que l'intelligence artificielle à code source ouvert et l'infrastructure publique numérique, entre autres méthodes et modèles commerciaux ; à promouvoir, à protéger et à préserver la diversité linguistique et culturelle, en tenant compte du multilinguisme dans leurs données d'apprentissage, en particulier pour les grands modèles de langage ; à adopter des mesures proactives pour lutter contre le racisme, la discrimination et d'autres formes de biais algorithmiques ; à permettre une concurrence loyale sur le marché de l'intelligence artificielle, à promouvoir un environnement propice à l'innovation et l'utilisation de biens publics numériques pour exploiter le potentiel de systèmes d'intelligence artificielle sûrs, sécurisés et dignes de confiance ;

OCDE – Rec. sur l’IA, 22 mai 2019



l’IA promet d’améliorer la prospérité et le bien-être des individus, de contribuer à une activité économique mondiale dynamique et durable, de stimuler l’innovation et la productivité, et d’aider à affronter les grands défis planétaires ; [...] la confiance est un déterminant essentiel de la transformation numérique ; que, bien qu’il soit difficile de prévoir la nature des futures applications de l’IA et leurs incidences, la confiance dans la fiabilité des systèmes d’IA est un facteur clé de la diffusion et de l’adoption de l’IA

OCDE, Recommandation sur l'IA (OECD/LEGAL/0449) (préambule)

- Première norme intergouvernementale sur l’IA
- Fournit la définition du système d'IA reprise par l'UE et le Conseil de l'Europe.
- Promeut une approche principielle
- *Soft law* largement repris (RIA, Principes du G20)

1.1. Croissance inclusive, développement durable et bien-être

Les parties prenantes devraient adopter de manière proactive une approche responsable en appui d’une IA digne de confiance afin de tendre vers des résultats bénéfiques pour les individus et la planète, tels que l’augmentation des capacités humaines et le développement de la créativité, l’inclusion des populations sous-représentées, la réduction des inégalités économiques, sociales, entre les sexes et autres, et la protection des milieux naturels, favorisant ainsi la croissance inclusive, le bien-être, le développement durable et la durabilité environnementale.

1.2. Respect de l’état de droit, des droits humains et des valeurs démocratiques, y compris de l’équité et de la vie privée

- a) Les acteurs de l’IA devraient respecter l’état de droit, les droits humains, les valeurs démocratiques et les valeurs centrées sur l’humain, tout au long du cycle de vie des systèmes d’IA. Cela comprend la non-discrimination et l’égalité, la liberté, la dignité, l’autonomie des individus, la protection de la vie privée et des données, la diversité, l’équité, la justice sociale, ainsi que les droits des travailleurs reconnus à l’échelle internationale. Cela recouvre aussi la lutte contre la désinformation et la désinformation amplifiées par l’IA, dans le respect de la liberté d’expression et des autres droits et libertés protégés par le droit international en vigueur.
- b) Pour ce faire, les acteurs de l’IA devraient instituer des garanties et des mécanismes, tels que la capacité d’intervention et de surveillance humaines, y compris pour faire face aux risques découlant d’utilisations à des fins autres que celles prévues ou d’une mauvaise utilisation délibérée ou involontaire, d’une manière qui soit adaptée au contexte et à l’état de l’art.

1.3. Transparence et explicabilité

Les acteurs de l’IA devraient s’engager à assurer la transparence et une divulgation responsable des informations liées aux systèmes d’IA. À cet effet, ils devraient fournir des informations pertinentes, adaptées au contexte et à l’état de l’art, afin :

- i. de favoriser une compréhension générale des systèmes d’IA, y compris de leurs capacités et de leurs limites,
- ii. d’informer les parties prenantes de leurs interactions avec les systèmes d’IA, y compris dans la sphère professionnelle,
- iii. de donner, lorsque cela est réalisable et utile, des informations claires et facilement compréhensibles sur les sources des données/entrées, les facteurs, les processus et/ou la logique ayant abouti à la prévision, au contenu, à la recommandation ou à la décision, pour permettre aux personnes concernées par un système d’IA d’en appréhender le résultat, et
- iv. de fournir des informations permettant aux personnes qui subissent les effets néfastes d’un système d’IA d’en contester les résultats.

1.4. Robustesse, sûreté et sécurité

- a) Les systèmes d’IA devraient être robustes, sûrs et sécurisés tout au long de leur cycle de vie, de sorte que, dans des conditions d’utilisation normales ou prévisibles, ou en cas d’utilisation abusive ou de conditions défavorables, ils soient à même de fonctionner convenablement, et ne fassent pas peser des risques de sûreté et/ou de sécurité démesurés.
- b) Des mécanismes devraient être en place, le cas échéant, afin de garantir que, dans l’éventualité où des systèmes d’IA risqueraient de causer des préjudices injustifiés ou présenteraient un comportement indésirable, ils puissent être neutralisés, réparés et/ou mis hors service en toute sécurité, en tant que de besoin.
- c) Des mécanismes devraient également être en place, lorsque cela est techniquement réalisable, pour renforcer l’intégrité de l’information, tout en veillant au respect de la liberté d’expression.

1.5. Responsabilité

- a) Les acteurs de l’IA devraient être responsables du bon fonctionnement des systèmes d’IA et du respect des principes exposés ci-dessus, selon leurs rôles, le contexte et l’état de l’art.
- b) Pour ce faire, les acteurs de l’IA devraient veiller à la traçabilité, notamment pour ce qui est des ensembles de données, des processus et des décisions prises au cours du cycle de vie des systèmes d’IA, afin de permettre l’analyse des résultats issus desdits systèmes d’IA et le traitement des demandes d’information, compte tenu du contexte et de l’état de l’art.
- c) Les acteurs de l’IA devraient, selon leurs rôles respectifs, le contexte et leur capacité à agir, appliquer de manière continue une approche systématique de la gestion du risque, à chaque phase du cycle de vie des systèmes d’IA, et assurer la conduite responsable des entreprises afin de gérer les risques y afférents, notamment, en tant que de besoin, par la coopération entre les différents acteurs de l’IA, les pourvoyeurs de connaissances et de ressources liées à l’IA, les utilisateurs de systèmes d’IA et les autres parties prenantes. Ces risques sont ceux liés aux biais préjudiciables, au non-

G7 – Le Processus d'Hiroshima

- Principes directeurs internationaux
- Code de conduite pour les organisations développant des systèmes d'IA avancés



Trois prolongements sectoriels

- Droit international humanitaire – l'IA de ciblage.
- Droit international au développement – la concentration mondiale de l'IA.
- Droit international des droits de l'homme – la désinformation et l'IA.

Droit international humanitaire



Les attaques seront strictement limitées aux objectifs militaires. [...] Ceux qui décident ou préparent une attaque doivent prendre toutes les précautions pratiquement possibles.

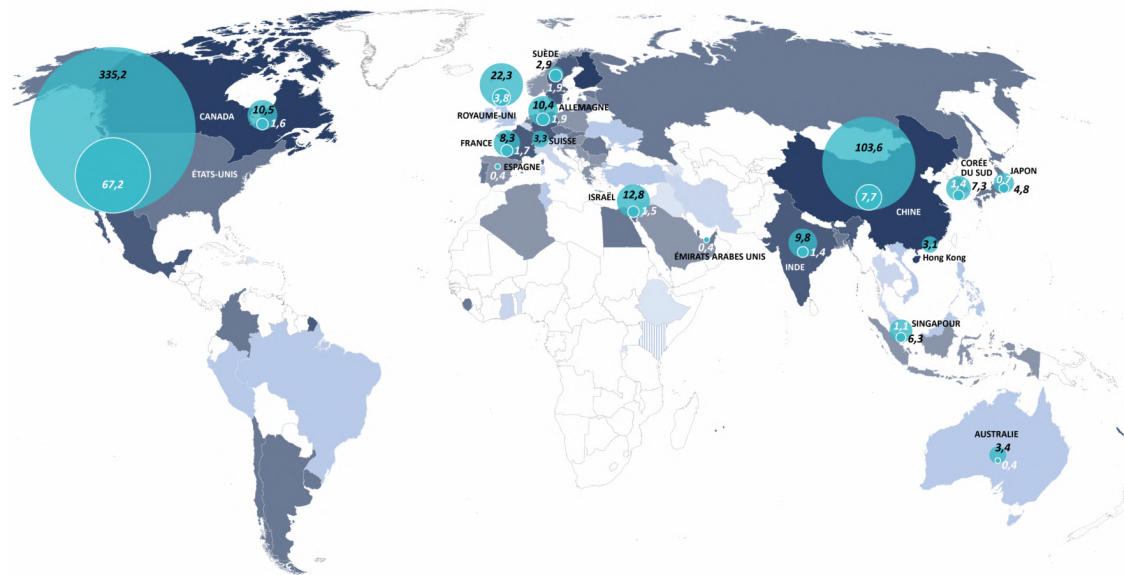
Protocole additionnel I aux Conventions de Genève (1977), art. 52, §2 et art. 57

- Principes cardinaux : distinction, proportionnalité, précautions dans toute attaque.
- L'IA de ciblage interroge l'« humain dans la boucle » et la responsabilité du commandement.
- Débats sur les systèmes d'armes létaux autonomes (SALA)

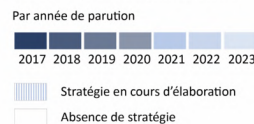
Droit au développement

- Pertinence d'une lecture sous l'angle du droit au développement
- La fracture algorithmique concentre les données, le calcul, l'énergie et les compétences au Nord, mais pas les ressources
- Pacte numérique Mondial (2024)

Panorama des puissances en intelligence artificielle (IA)



Les stratégies nationales en IA

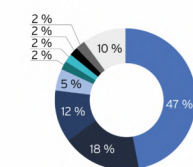


Principaux pays en matière d'investissements privés dans l'IA



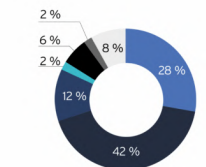
Origine des principaux chercheur-euse-s en IA, 2022

En prenant en compte leur formation universitaire de premier cycle et les 20 % de chercheur-euse-s en IA les plus performants



Lieu de travail des principaux chercheur-euse-s en IA, 2022

En prenant en compte les 20 % de chercheur-euse-s en IA les plus performants



Sources : Artificial Intelligence Index Report 2024, Stanford University ; Statista (Cloudscene) ; MacroPolio

Droit international des droits de l'homme – l'exemple de la désinformation

- L'IA au service de la désinformation
- La désinformation produite par l'IA

