

COMPRENDRE, DÉBATTRE,
PROPOSER & AGIR :
L'IA DE DEMAIN SE DÉCIDE
ICI ET MAINTENANT !

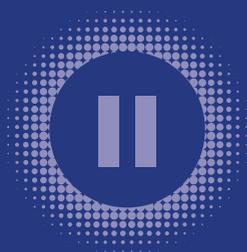


REPRENDRE LE CONTROLE

FORUM DES SOLUTIONS
POUR UNE IA COMPATIBLE
AVEC L'HUMANITÉ

8 FÉVRIER 2025
LEARNING PLANET INSTITUTE, PARIS

 PAUSE IA



SOMMAIRE

- P.3 CONTEXTE ET VISION
- P.5 PROGRAMME DE LA JOURNÉE
- P.7 LES INTERVENANTS
- P.13 À PROPOS DE PAUSE IA
- P.15 INFOS PRATIQUES PRESSE

CONTEXTE ET VISION

UN SOMMET NÉ POUR RÉPONDRE À UN DÉFI EXISTENTIEL

Le Sommet pour la Sécurité de l'IA, lancé par le gouvernement britannique à Bletchley en novembre 2023, a émergé en réaction aux risques sans précédent liés au développement rapide de l'intelligence artificielle. Dans un contexte où des voix parmi les plus reconnues du secteur ont alerté sur la dimension potentiellement « existentielle » des menaces que représente l'IA, ce premier Sommet fut l'occasion pour de nombreux États, dont la France, de signer la Déclaration de Bletchley. Ce texte historique reconnaît explicitement le potentiel de dommages catastrophiques qu'offrent les modèles d'IA avancés : **biosécurité, cybersécurité et désinformation**, entre autres.

UN CONTEXTE DE SURENCHÈRE TECHNOLOGIQUE ET DE VULNÉRABILITÉS CROISSANTES

La fin de l'année 2024 a été marquée par une série d'annonces spectaculaires sur les progrès de l'IA, notamment de la part de OpenAI, qui a présenté son nouveau modèle de raisonnement O3, pulvérisant des seuils historiques sur plusieurs benchmarks. **Étonnamment, ces percées ont reçu peu de couverture médiatique, alors même que les recherches en sécurité de l'IA multiplient les signaux d'alarme quant à l'émergence de capacités de plus en plus dangereuses.** Simultanément, la puissance grandissante des réseaux sociaux fait peser un risque inédit sur la stabilité des démocraties et la fiabilité de l'information, tandis que la cybercriminalité connaît une accélération exponentielle.

DU SOMMET POUR LA SÉCURITÉ... AU SOMMET POUR L'ACTION

Dans ce climat d'urgence, la France devait organiser en 2025 une suite logique au Sommet de Bletchley. Cependant, la version française a été baptisée « **Sommet pour l'Action sur l'IA** » et a réduit la thématique de la sécurité au second plan, au profit d'autres axes. Cette évolution a, selon plusieurs observateurs et notre propre contre-expertise, été favorisée par un lobbying important de grands acteurs technologiques. Le terme "**sécurité**" n'apparaît d'ailleurs plus parmi les axes de travail officiels, **soulevant des inquiétudes quant à la pleine reconnaissance des menaces que l'IA fait peser sur nos sociétés.**

UNE INITIATIVE CITOYENNE POUR REPLACER LA SÉCURITÉ AU CŒUR DES DÉBATS

Face à cette omission, de nombreuses organisations de la société civile ont décidé de prendre l'initiative. Notre Forum « *Reprendre le Contrôle : Forum des solutions pour une IA compatible avec l'humanité* » répond à ce besoin : **rassembler experts, chercheurs, décideurs et citoyens afin d'examiner les risques réels posés par l'IA, et d'identifier des solutions concrètes pour y faire face.** C'est une démarche résolument ascendante (bottom-up), qui cherche à remettre la sécurité et la gouvernance responsable de l'IA au centre de l'agenda.

OBJECTIF : INFLUENCER LE SOMMET DANS UN ESPRIT CONSTRUCTIF

Au-delà des débats et tables rondes qui se tiendront le 8 février 2025, notre ambition est de produire dès le lendemain **un débrief collectif qui sera diffusé auprès des participants du Sommet pour l'Action sur l'IA.** Ce document, fruit des réflexions de la société civile, a pour vocation d'éclairer les décideurs et médias sur la nécessité de prendre en compte les enjeux sécuritaires et éthiques – pourtant fondamentaux – dans la stratégie globale de régulation de l'IA.

UN APPEL AUX MÉDIAS

Nous en appelons aux journalistes et médias pour relayer ces réflexions et contribuer à faire émerger une voix citoyenne forte, portée par l'intérêt général. Notre Forum souhaite apporter un éclairage complémentaire, voire correctif, à une conversation qui risque de se faire sans la mise en avant suffisante des questions de sécurité, malgré l'importance vitale de ces dernières.

«IL N'Y A QUE
DEUX MANIÈRES
DE RÉAGIR,
À UN PHÉNOMÈNE
EXPONENTIEL:
TROP TÔT
OU
TROP TARD.»

CONNOR LEAHY
CONTROL AI

PROGRAMME DE LA JOURNÉE

MATINÉE

- 9h30 – 10h00

ACCUEIL DES PARTICIPANTS ET OUVERTURE

- 10h00 – 11h00

TABLE RONDE #1

DE L'ANTICIPATION À LA RÉALITÉ : ÉVALUER ET PRIORISER LES RISQUES DE L'IA

Alors que nous développons des systèmes d'IA toujours plus puissants, de nombreux risques autrefois théoriques sont devenus réalité. Comment orienter le développement de ces technologies pour éviter les écueils actuels tout en prévenant l'émergence de dangers potentiellement plus graves ?

Modération

Raja Chatila

Avec

**Marc Faddoul, Jeremy Perret,
Mathilde Cerioli, Arthur Grimonpont**

- 11h00 – 11h30

PAUSE CAFÉ & NETWORKING

- 11h30 – 13h00

KEYNOTES :

- **MAXIME FOURNES**

Président de **Pause IA**

- **LÊ NGUYÊN HOANG**

PDG de **Calicarpa** et Président
de **Tournesol**

- 13h00 – 14h00

PAUSE DÉJEUNER & VISITE DES STANDS

APRÈS-MIDI

- 14h00 – 15h30

TABLE RONDE #2

SÉCURISER L'AVENIR DE L'IA, APPROCHES TECHNIQUES ET RÉGLEMENTAIRES

Face aux défis croissants de la sécurité de l'IA, cette table ronde explorera comment les solutions techniques et les mesures de gouvernance peuvent se compléter pour créer un cadre efficace de développement sûr de l'IA.

Modération

Nasri Badaoui

Avec

**Flavien Chervet,
Charbel-Raphael Segerie,
Anne-Sophie Simpere,
Lou Welgryn**

- 15h30 – 16h00

PAUSE CAFÉ & NETWORKING

- 16h00 – 17h30

PRÉSENTATIONS FLASH ET CLÔTURE

Session de présentations courtes mettant en lumière des initiatives innovantes dans le domaine de la sécurité de l'IA.

EN CONTINU

Des stands interactifs seront accessibles tout au long de la journée, présentant différentes organisations et projets œuvrant pour une IA responsable.

LES INTERVENANTS

MATHILDE CERIOLI

INTERVENANTE CHEZ EVERYONE AI



Docteure en neurosciences, Mathilde Cerioli s'est spécialisée dans le développement de l'enfant et l'intégration de l'IA pour concevoir des environnements numériques sûrs. Entrepreneure et consultante, elle privilégie une approche ludique et créative, propice à l'innovation et à la résolution de défis technologiques et sociétaux. **Grâce à sa double expertise en neurosciences et entrepreneuriat, elle encourage les individus et les organisations à miser sur l'imagination pour construire des solutions transformatrices et positives.**

 www.linkedin.com/in/mathilde-neugnot-cerioli/

LOU WELGRYN

INTERVENANTE CHEZ DATA FOR GOOD



Lou Welgryn est co-présidente de **Data For Good**, association qui réunit plus de 4000 spécialistes en technologies numériques pour soutenir des projets à fort impact social et environnemental. Elle est également **Carbon Data Analyst** chez **Carbone4 Finance**, où elle analyse les données d'émissions de gaz à effet de serre pour guider les investissements vers des solutions bas carbone. Ancienne élève de l'ESSEC, elle a quitté la finance traditionnelle pour se consacrer à la transition écologique, **incarnant une vision humaniste de la data mise au service du bien commun.**

 www.linkedin.com/in/lou-welgryn-460434b0/



LÊ NGUYỄN HOANG

CHERCHEUR, PRÉSIDENT DE L'ASSOCIATION Tournesol

Polytechnicien, docteur en mathématiques, Lê Nguyễn Hoàng a été chercheur au MIT et à l'EPFL. **Il est PDG cofondateur de Calicarpa**, créateur de la chaîne YouTube **Science4All**, président de l'association **Tournesol** et membre du conseil d'éthique d'Orange. Il est également coauteur de plusieurs ouvrages, dont *La Dictature des Algorithmes*, où il envisage un numérique sécurisé et démocratique.

www.linkedin.com/in/l%C3%AA-nguy%C3%AAn-hoang/



OTTO BARTEN

CHERCHEUR AU X-RISK OBSERVATORY

Otto Barten est ingénieur spécialisé en énergie durable, data scientist et entrepreneur. Constatant que certains risques existentiels surpassent même les enjeux climatiques, il a fondé l'**Existential Risk Observatory** à Amsterdam, une organisation dédiée à la sensibilisation sur les risques d'extinction humaine. Diplômé en physique théorique et en technologies énergétiques durables, il a mené des projets innovants dans les domaines de l'énergie et de la durabilité.

www.linkedin.com/in/ottobarten/



FLAVIEN CHERVET

EXPERT

Flavien Chervet est un penseur multidisciplinaire, alliant philosophie, arts, design, sciences et technologies. **Co-fondateur et Directeur Innovation d'IRIIG** (grande école française de l'innovation), il est également vice-président de Lyon-iS-Ai, un do-tank dédié à l'éthique et à la démystification de l'intelligence artificielle. Auteur des ouvrages *Hypercréation* et *Hyperprompt*, il étudie depuis 2014 les interactions entre IA et créativité humaine, **tout en promouvant des usages éthiques et éclairés des technologies.**

www.linkedin.com/in/flavienchervet/



ARTHUR GRIMONPONT

INGÉNIEUR ET CONSULTANT EN TRANSITION ÉCOLOGIQUE

Arthur Grimonpont est un ingénieur et consultant spécialisé dans la transition de nos sociétés face aux bouleversements climatiques et écologiques. Cofondateur de l'association **Les Greniers d'Abondance**, il promeut des politiques de résilience alimentaire adaptées à un monde à +3°C. Engagé au sein de

Reporters sans Frontières, il y dirige le desk **IA & enjeux globaux**. Il a en particulier été rapporteur de la **Charte de Paris sur l'IA et le journalisme**, pilotée par la lauréate du Prix Nobel de la paix Maria Ressa. Auteur de l'essai *Algocratie* (Actes Sud, 2022), il y explore **le rôle des réseaux sociaux et de leurs algorithmes de recommandation dans la diffusion de l'information et la formation des opinions**.



MAXIME FOURNES

FONDATEUR DE PAUSE IA

Maxime Fournes est le président de l'association **Pause IA**, le chapitre français de **Pause AI**, dédiée à sensibiliser aux risques liés au développement incontrôlé de l'intelligence artificielle.

Ancien **expert en deep learning et machine learning**, il a quitté le domaine pour tirer la sonnette d'alarme sur les menaces existentielles que pose cette technologie. Diplômé de Centrale Lyon et de l'Université de Cambridge, **il combine une solide expertise scientifique avec un engagement passionné pour promouvoir une IA responsable et sécurisée**.



MARC FADDOUL

DIRECTEUR ET COFONDATEUR D'AI FORENSICS

Marc Faddoul est un ingénieur transdisciplinaire spécialisé dans les **systèmes de recommandation et l'audit algorithmique**. Co-fondateur et directeur d'**AI Forensics**, une ONG européenne dédiée à l'investigation des algorithmes opaques

et influents, il œuvre à tenir les grandes plateformes technologiques responsables de leurs impacts sociétaux. Formé à Télécom Paris et UC Berkeley, Marc a acquis une solide expertise chez Facebook AI et dans le milieu des start-ups, tout en collaborant avec des institutions européennes et des médias internationaux sur les enjeux éthiques et sociétaux de l'IA.



JÉRÉMY PERRET

EXPERT

Jérémy Perret est docteur en intelligence artificielle, spécialisé dans le traitement automatique du langage naturel.

Formé à l'ENSEEIH (informatique et mathématiques appliquées) et titulaire d'un doctorat de l'IRIT, il a développé des modèles prédictifs pour des applications innovantes en assistance professionnelle. Il travaille aujourd'hui comme producteur indépendant chez **Suboptimal IA**, où il poursuit ses recherches et ses projets à **la croisée de la science et de l'innovation technologique.**

www.linkedin.com/in/j%C3%A9r%C3%A9my-perret-b22a72b9/



ADAM SHIMI

RESPONSABLE DE L'ITÉRATION CHEZ CONTROLAI

Après une formation d'ingénieur à l'ENSEEIH et un doctorat à l'IRIT, il s'est orienté vers la recherche en sûreté de l'IA, travaillant d'abord en indépendant puis chez **Conjecture**, une start-up dédiée aux risques catastrophiques liés à l'IA. Il s'y est impliqué dans la recherche, l'ingénierie, la rédaction technique et la conception de produits, avant de se consacrer à la diffusion des risques et des solutions, tant techniques que civiques. Aujourd'hui, il est responsable de l'itération chez **ControlAI**, où il pilote notamment **The Compendium** (une perspective globale sur l'IA moderne et ses risques) et **A Narrow Path** (plan concret et détaillé pour gérer les risques catastrophiques de l'IA tout en préservant l'innovation).

www.linkedin.com/in/adam-shimi-23b27a125/



STÉPHANE AMARSY

EXPERT

Stéphane Amarsy est un expert en IA, innovation et prospective, passionné par la transformation des organisations face aux défis de demain. Il est PDG de **The Next Mind**, qu'il décrit comme un accompagnateur pour aider les entreprises à trouver leur «cœur anti-fragile» et faire face aux bouleversements technologiques, économiques et sociétaux. Auteur et conférencier, il développe une approche holistique où technologies et dynamiques humaines se conjuguent pour bâtir des solutions durables, porté par sa devise : *« S'occuper du changement avant qu'il ne s'occupe de vous ! »*

www.linkedin.com/in/stephane-amarsy/



CHARBEL-RAPHAEL SEGERIE

DIRECTEUR EXÉCUTIF DU CESIA (CENTRE POUR LA SÉCURITÉ DE L'IA)

Charbel-Raphael Segerie est le **Directeur exécutif du Centre pour la Sécurité de l'IA (CeSIA)** à Paris, où il **pilote la recherche et la formation en matière de sûreté de l'IA**. Il enseigne un cours spécialisé à l'**École Normale Supérieure**, en cours de formalisation dans un manuel dédié. Ses travaux portent sur l'identification des risques émergents de l'IA, l'amélioration des méthodes de sûreté actuelles (RLHF, interprétabilité) et le développement d'approches *safe-by-design*. Il participe aussi à l'effort européen d'évaluation de l'IA et à la rédaction du **Code de Pratique pour l'IA généraliste** auprès du EU AI Office. Charbel a également fondé **ML4Good**, un bootcamp d'une dizaine de jours visant à former des chercheurs en IA Safety.

www.linkedin.com/in/charbel-raphael-segerie/



RAJA CHATILA

PROFESSEUR ÉMÉRITE À SORBONNE UNIVERSITÉ

Raja Chatila est **professeur émérite d'Intelligence Artificielle, de Robotique et d'éthique à Sorbonne Université**, où il a dirigé l'Institut des Systèmes Intelligents et de Robotique (ISIR) et le laboratoire d'excellence SMART. Auteur de plus de 180 publications, Fellow de l'IEEE, il a reçu le **IEEE "Robotics and Automation Pioneer Award."** Co-président du groupe **"IA responsable"** du **Partenariat Mondial sur l'IA (GPAI)**, il a également été membre du **groupe d'experts de haut niveau sur l'IA de la Commission européenne**.



À PROPOS DE PAUSE IA

APPRÉHENDER LES RISQUES LIÉS À L'INTELLIGENCE ARTIFICIELLE

Pause IA est une association à but non lucratif fondée en 2024, dont la mission est de sensibiliser et de mobiliser le grand public, les décideurs politiques et la communauté scientifique face aux risques majeurs que l'IA va faire peser sur l'humanité. La branche française de Pause AI a été créée par Maxime Fournes, ancien chercheur en deep learning passé par Cambridge pour ses études de mathématiques, puis chief data scientist à Londres. Convaincu que le développement d'une IA non encadrée menace notre avenir collectif, il ouvre le chapitre français de Pause AI pour défendre l'idée en France d'un **moratoire international** dans la course à l'Intelligence Artificielle Générale (IAG), tout en promouvant une approche globale et responsable du développement de l'IA.

NOS ACTIONS CONCRÈTES

→ Contre-expertise du rapport gouvernemental sur l'IA (mars 2023)

Nous avons publié un **contre-rapport**, soutenu par une vingtaine d'experts et organisations indépendantes, proposant une analyse approfondie du rapport officiel du Comité gouvernemental sur l'IA. Cette étude met en lumière des angles morts significatifs dans l'évaluation des risques, et souligne l'influence qu'ont pu avoir les acteurs industriels majeurs dans l'orientation des conclusions du rapport initial.

 <https://contre-rapport-ia.fr/>

→ Analyse des dangers de l'IA

Sur notre site, nous présentons les **dangers économiques, matériels et sociétaux** que l'IA fait peser sur nos sociétés, ainsi que des pistes d'action concrètes pour minimiser ces menaces.

→ Dialogue institutionnel et sensibilisation du public

Nous multiplions les rencontres avec des parlementaires, acteurs publics et privés pour encourager une régulation intelligente et ambitieuse. En parallèle, nous organisons **conférences, publications, tribunes et manifestations** afin d'informer et d'alerter l'ensemble de la société civile.

UNE VISION INTERNATIONALE ET RESPONSABLE

Face à l'accélération du développement de l'Intelligence Artificielle, Pause IA milite pour la mise en place d'un **moratoire international** visant à limiter la course technologique vers l'IAG et, à terme, la **Superintelligence**. Nous défendons l'idée qu'un cadre mondial est indispensable pour que l'IA reste bénéfique à l'humanité. Le **Sommet pour l'Action sur l'IA** (10-11 février 2025) constitue une occasion historique pour consolider cette gouvernance globale.

NOTRE ÉVÉNEMENT : **LE FORUM DES SOLUTIONS POUR UNE IA COMPATIBLE AVEC L'HUMANITÉ**

Le 8 février 2025, Pause IA organise à Paris une journée de conférences, tables rondes et ateliers, qui s'inscrit dans la continuité de notre ambition :

→ **Réunir les principaux acteurs de la sécurité de l'IA en France — chercheurs, décideurs, entrepreneurs, militants —** pour débattre des crises possibles et proposer des solutions concrètes.

→ **Créer de l'intelligence collective** : Favoriser les échanges et la confrontation de points de vue, afin d'éclairer les choix politiques et technologiques.

→ **Faire connaître l'écosystème** : Présenter au grand public les femmes et les hommes qui œuvrent déjà pour la sécurisation de l'IA, à quelques jours seulement du Sommet pour l'Action sur l'IA.

Cette proximité avec le Sommet offre un momentum exceptionnel : la France disposera d'une tribune internationale pour faire entendre un message fort, collectif et expert, rappelant que seule une gouvernance internationale ferme de l'IA peuvent préserver notre avenir commun. nécessaire pour préserver notre avenir commun.

INFOS PRATIQUES PRESSE

Afin de faciliter la couverture médiatique du **Forum des Solutions pour une IA Compatible avec l'Humanité**, vous trouverez ci-dessous toutes les informations nécessaires pour votre accréditation et vos demandes d'interviews.

CONTACTS PRESSE

Ludwig Iguedjal – Responsable Presse
Email : ludwig@pauseia.fr
Téléphone : **07.68.25.02.14**

ACCREDITATION ET BADGES PRESSE

- **Demande d'accréditation** : Merci d'envoyer un email à ludwig@pauseia.fr en indiquant vos nom, prénom, média représenté et coordonnées.
- **Confirmation** : Vous recevrez sous 48h un email de confirmation avec les modalités de retrait de votre badge presse le jour de l'événement.
- **Badge** : Un espace dédié sera prévu à l'accueil pour récupérer votre badge. Veuillez vous munir d'une pièce d'identité et de votre carte presse, si vous en disposez.

PROCÉDURE D'INTERVIEW ET DE REPORTAGE

- **Demande d'interview** : Pour toute demande d'interview (intervenants, organisateurs, etc.), contactez-nous à l'adresse ludwig@pauseia.fr. Vous pouvez préciser vos centres d'intérêt (thématique, intervenant, etc.) pour faciliter l'organisation.

- **Planification** : Les créneaux d'interview pourront être réservés en amont ou le jour J, selon les disponibilités.
- **Reportage sur place** : Si vous prévoyez de filmer ou de réaliser des reportages audio/vidéo, merci de le signaler à l'avance afin que nous puissions vous réserver un espace adapté.

RESSOURCES TÉLÉCHARGEABLES

- **Visuels et Logos** : Une sélection de logos et de photos haute résolution (intervenants, organisateurs, images d'illustration) est disponible sur notre espace presse en ligne :

 https://drive.google.com/drive/u/0/folders/1oaB_X17quWi9N18nrPX9FPDwlf-duIQNo

MENTIONS LÉGALES ET RÈGLES SPÉCIFIQUES

- **Droits d'image** : Toute utilisation d'images, d'extraits sonores ou de vidéos doit s'effectuer dans le respect de la législation en vigueur et après validation par l'équipe presse, si nécessaire.
- **Validation** : Toute demande de reportage, d'interview ou de diffusion d'images complémentaires doit être adressée à ludwig@pauseia.fr et validée par l'organisation.

Si vous avez la moindre question, n'hésitez pas à nous contacter directement par email ou par téléphone. Nous restons à votre disposition pour vous aider à préparer votre venue et faciliter votre travail de couverture médiatique.



REPRENDRE LE CONTROLE

FORUM DES SOLUTIONS
POUR UNE IA COMPATIBLE
AVEC L'HUMANITÉ

8 FÉVRIER 2025

LEARNING PLANET INSTITUTE

8 BIS RUE CHARLES V, 75004 PARIS

M° SULLY - MORLAND / SAINT-PAUL / BASTILLE

CONTACTS PRESSE

LUDWIG IGUEDJAL – RESPONSABLE PRESSE

EMAIL : LUDWIG@PAUSEIA.FR

TÉLÉPHONE : 07.68.25.02.14