Chers membres du réseau de l'ancienne FSFA, chers intéressés,

Les principaux laboratoires américains ont demandé que le développement de l'IA soit ralenti, car les mesures de sécurité ne parviennent pas à suivre le rythme exponentiel de la recherche en IA. Les réactions à cet appel à alerte ont été diverses aux États-Unis, en Chine et en Europe. Il existe toutefois un consensus sur la nécessité urgente de standards de sécurité convenus au niveau international. En raison d'une perception différente du problème et d'intérêts divergents, les deux superpuissances de l'IA, les États-Unis et la Chine, ont de la peine à parvenir rapidement à un consensus. Une médiation s'impose. Cela ouvre des perspectives pour l'Europe, mais aussi pour la Suisse.

**États-Unis : des mesures d'autocontrôle volontaires et un nouveau terme**
Pour Trump, mais aussi pour d'autres personnalités telles que le CEO de Nvidia, Jensen Huang, la course à l'IA avec la Chine ne permet aucun ralentissement. Il suffirait que les laboratoires d'IA mettent en place des mesures d'autocontrôle volontaires, lesquelles ont été définies le 29 septembre 2026 dans une déclaration sur la « Super Intelligence ». Les directives qui y figuraient étaient toutefois moins strictes que les exigences auxquelles les laboratoires d'IA basés dans la Silicon Valley se conformaient déjà en raison de la loi californienne Transparency in Frontier AI Act (SB 53). La seule véritable nouveauté était en réalité le décret présidentiel de Trump, selon lequel le terme « Artificial Intelligence » devait désormais être remplacé par « Super Intelligence ». Le gouverneur californien a réagi sans tarder en ordonnant que le nom « Artificial Intelligence » reste utilisé en Californie. En tant que quatrième économie mondiale, cet État fédéré peut se permettre de faire opposition, et OpenAI n'est pas obligé de changer son nom en OpenSI.

**La Chine a une vision totalement différente de la question de la sécurité de l'IA**
L'appel au ralentissement est perçu en Chine comme une manœuvre visant à freiner le développement de l'IA, et donc comme un retour à la guerre froide. La Chine n'est pas en tête dans le domaine de la recherche et a déjà promulgué des directives en matière de sécurité de l'IA. Les préoccupations des dirigeants chinois portent avant tout sur les valeurs véhiculées par les modèles d'IA, ainsi que sur la stabilité sociale et la sécurité nationale. Tous les modèles d'IA sont donc soumis à une censure étatique. La Chine mise principalement sur les modèles open source et leur exportation via des coopérations en matière d'IA. Ces outils sont téléchargeables, peu coûteux et sont donc utilisés dans le monde entier, y compris en Occident.

**Une alternative européenne**
L'Europe est en train d'emprunter une troisième voie, grâce à des modèles d'IA qui allient souveraineté nationale et valeurs démocratiques. La start-up française Mistral vient de lancer Large 4, « Le chonk », un grand modèle d'IA qui surpasse ses concurrents américains et chinois. Avec ses modèles conformes aux normes européennes, qu'ils soient propriétaires ou open source, Mistral propose des alternatives européennes qui s'imposent dans un nombre croissant de pays.

**L'Europe s'engage en faveur de la recherche sur la sécurité**
Ces dernières années, la recherche de pointe en IA s'est déplacée des grandes écoles et universités vers les laboratoires privés spécialisés dans l'IA, ces derniers disposant de ressources bien plus importantes. Avec leur initiative « A Call for Open Science in AI Safety », lancée par l'ETHZ et l'EPFL, les hautes écoles ont indiqué qu'elles soutiendraient les modèles d'IA par le biais de leur recherche en matière de sécurité, à condition d'avoir accès aux codes, aux processus d'entraînement, aux mesures de sécurité et aux données d'évaluation.

Peu après, plus de 20 pays, principalement européens, ainsi que la présidente de la Commission européenne, ont signé une initiative internationale intitulée « A Call for Control of Frontier AI Models ». Celle-ci exige également un accès aux « Frontier Labs », ces laboratoires à la pointe de la recherche. Elle réclame en outre des normes de sécurité communes et la création d'une autorité de contrôle internationale.

À peine une semaine plus tard, l'argumentation scientifique de cette initiative politique a été présentée dans un document consacré à l'explosion des connaissances à venir. Ce document préconise, comme mesure de sécurité nécessaire, la présence d'experts externes indépendants, intégrés au sein des laboratoires d'IA. Parmi les auteurs figurent également des responsables des deux laboratoires Anthropic et OpenAI.

**Une opportunité pour la Suisse** Le cinquième sommet international sur l'IA se tiendra les 21 et 22 juin 2027 à Genève. Les travaux préparatoires avancent déjà à grands pas, et il y a de fortes chances que la Suisse parvienne à tracer des voies consensuelles pour un développement futur plus sûr de l'IA, dans l'intérêt de tous.

Avec nos salutations les meilleures,
Pour le réseau de l'ancienne FSFA : Hanna Muralt Müller 2.11.2026

***Si vous ne souhaitez plus recevoir cet e-mail, veuillez me contacter : info@muralt-mueller.ch.***