



Un an après sa signature, le continuum public-privé Scaleway / GENCI / CNRS porte ses premiers fruits : la RATP déploie “Sillon” son premier modèle IA souverain en production

Vivatech, mercredi 18 juin 2026

Annoncé à VivaTech 2025, le partenariat entre Scaleway, GENCI et le CNRS affiche des premiers résultats concrets : une interopérabilité opérationnelle entre infrastructures publiques et privées, une dizaine de projets accompagnés, et une première démonstration industrielle avec la RATP et la startup Pleias.

Le pari de l'interopérabilité, tenu

Il y a un an, à Vivatech 2025, Scaleway, le CNRS et GENCI annonçaient la création d'un continuum public/privé souverain d'infrastructures de calcul pour accélérer le développement et les usages de l'IA. L'ambition : créer des passerelles techniques entre les supercalculateurs publics — au premier rang desquels Jean Zay de GENCI, hébergé et opéré par l'Institut du développement et des ressources en informatique scientifique du CNRS (IDRIS) — et les clusters IA privés de Scaleway, afin de permettre aux chercheurs et aux entreprises de faire circuler leurs travaux d'IA, de l'apprentissage au fine-tuning puis à l'inférence à grande échelle, sans rupture technique.

Achevés au printemps 2026, ces travaux ont déjà permis à une dizaine de projets de bénéficier de cette fluidité d'usage. Ils démontrent qu'un modèle de gouvernance partagée entre secteur public et acteurs privés souverains est non seulement possible, mais désormais opérationnel.

La RATP, premier cas d'usage industriel du continuum

Parmi les projets accompagnés, celui de la RATP, développé avec la startup Pleias, est particulièrement emblématique. SILLON, premier modèle open source de la RATP, a été entraîné sur les moyens de calcul de GENCI et grâce à l'équipe support du CNRS à l'IDRIS dans le cadre du projet AI Factory France, puis déployé en inférence à grande échelle sur l'infrastructure souveraine privée de Scaleway, au service des agents de la RATP.

Sillon est un Small Language Model souverain (600 millions de paramètres) conçu par la RATP et Pleias pour raisonner sur les messages des voyageurs, notamment réseaux sociaux et service clientèle. Outil d'aide à la décision, il détecte et mesure en temps quasi réel les signes de détresse ou d'urgence. Son raisonnement, affiché aux opérationnels de supervision du réseau, les aide à interpréter efficacement les afflux de messages lors d'incidents.

Entraîné sur des données 100 % synthétiques et sur le supercalculateur Jean Zay (GENCI/CNRS), Sillon est jusqu'à 1 000 fois plus petit que les modèles généralistes tout en étant plus précis sur la classification. Son empreinte carbone est minimale : entraînement et dix ans d'exploitation n'équivalent qu'à quelques minutes de fonctionnement d'une ligne de métro.

Avec Pleias et le consortium CNRS-GENCI-Scaleway, la RATP fait une démonstration de Sillon et illustre le potentiel des modèles de raisonnement transparents issus de l'expertise métier d'une organisation.

« Au sein de la RATP, l'innovation s'attache à transformer, seul ou en partenariat, des avancées technologiques en solutions concrètes pour nos métiers. Notre objectif est d'être au bon niveau de technologie, là où elle apporte un gain réel pour les voyageurs, les agents, et les villes. Avec SILLON, nous démontrons ce savoir-faire avec des partenaires, pour une technologie utile, maîtrisée et pleinement intégrée aux exigences d'un grand opérateur de transport public. » déclare **Gilles Tauzin, directeur de l'innovation Groupe RATP**

Au-delà de l'interopérabilité : gestion du débordement, apprentissage fédéré et veille technologique

Le partenariat a également produit des résultats sur deux autres axes.

D'une part, les partenaires ont mené des travaux d'évaluation pour mettre en place des mécanismes de débordement afin d'utiliser les capacités GPU privées de Scaleway lorsque les ressources publiques sont saturées, dans un cadre de confiance et de gouvernance partagée. Ils ont en particulier travaillé autour de l'évaluation d'outils d'apprentissage fédéré, (NVIDIA FLARE, DeepMind DiLoCo) permettant la mise en commun de leurs puissances de calcul respectives pour entraîner des modèles IA de frontière tout en gardant localement les données, un enjeu majeur pour certains secteurs industriels ou pour la santé.

D'autre part, l'objectif annoncé de la mise en place d'un laboratoire commun, dédié à l'optimisation des charges IA et à la réduction de l'empreinte carbone du calcul intensif, s'est élargi à une veille technologique approfondie, en particulier autour de solutions d'inférence européennes.

Un maillon du futur écosystème européen de l'IA

Ces résultats s'inscrivent dans une dynamique plus large. Le projet AI Factory France, piloté par GENCI et labellisé par EuroHPC, fédère 16 partenaires publics et privés ainsi qu'un cercle de 40 partenaires associés. Il constitue l'un des 19 nœuds du réseau européen d'AI Factories, qui devra lui-même s'articuler avec les futures AI GigaFactories — infrastructures de plus de 100 000 GPUs — attendues à partir de 2027.

Le continuum Scaleway / GENCI / CNRS préfigure ainsi, à l'échelle française, le lien indispensable entre ces deux niveaux d'infrastructure que l'Europe devra tisser pour peser dans la compétition mondiale en matière d'IA.

« Je me félicite que le partenariat que nous avons noué l'année dernière produise aujourd'hui des résultats concrets. Nous souhaitons renforcer ces liens entre la recherche publique et le

monde industriel afin de rester aux côtés de l'excellence scientifique européenne et de porter l'innovation jusqu'au marché. », déclare **Damien Lucas, Directeur Général de Scaleway.**

“L'intelligence artificielle est devenue un enjeu de souveraineté et de compétitivité majeur. Avec la mise en place d'un continuum entre les infrastructures publiques de GENCI et privées de Scaleway, la France se dote d'un modèle original qui permet d'accélérer l'innovation tout en garantissant la maîtrise de ses technologies et de ses données. Les premiers résultats obtenus notamment par la RATP associée à la startup Pleias montrent que cette ambition est déjà une réalité.” déclare **Michaël Krajecki, PDG de GENCI**

« Le CNRS est fier de compter sur les équipes de l'IDRIS et son réseau d'ingénieurs et ingénieures en IA, financé dans le cadre du Programme national de recherche en intelligence artificielle (PNRIA). Centre de calcul national, l'IDRIS met au cœur de sa stratégie le service aux équipes utilisatrices et l'écoresponsabilité. Les ingénieurs et ingénieures en IA, répartis dans 5 laboratoires en France, offrent un accompagnement de qualité aux scientifiques qui portent leurs projets sur le supercalculateur Jean Zay. Ils optimisent l'exploitation des ressources de calcul au service du monde académique et des partenariats avec les acteurs socio-économiques. L'IDRIS et l'ensemble des équipes support sont un levier de la souveraineté en matière de recherche en intelligence artificielle », déclare **Mehdi Gmar, directeur général délégué à l'innovation du CNRS.**

À propos de GENCI

Créée par les pouvoirs publics en 2007, GENCI (Grand Équipement National de Calcul Intensif) est une grande infrastructure de recherche, opérateur public visant à démocratiser l'usage de la simulation numérique par le calcul haute performance associé à l'usage de l'intelligence artificielle et du calcul quantique pour soutenir la compétitivité scientifique et industrielle française.

GENCI poursuit trois missions :

- Mettre en œuvre la stratégie nationale d'équipement en moyens de calcul intensif, stockage et de traitement de données massives associé aux technologies de l'IA (Intelligence Artificielle) et aux futures technologies de calcul quantique au bénéfice de la recherche scientifique ouverte française et Européenne en s'appuyant sur les trois centres nationaux de calcul ;
- Soutenir la réalisation d'un écosystème intégré à l'échelle nationale et européenne
- Promouvoir la simulation numérique par le calcul intensif auprès des équipes de la recherche ouverte académique et industrielle

GENCI est une société civile détenue à 49 % par l'État représenté par le Ministère en charge de l'Enseignement Supérieur, de la Recherche et de l'Espace, 20 % par le CEA, 20 % par le CNRS, 10 % par les Universités représentées par France Universités et 1 % par Inria.



GENCI est le coordinateur du projet AI Factory France aux côtés d'Inria, CNRS, CEA, CINES, AMIAD, HubFranceIA, les 9 IA Clusters, l'Université de Reims Champagne Ardennes, French Tech, Station F et France Universités. Financée par EuroHPC et faisant partie d'un réseau de 19 AI Factories en Europe, AI Factory France offre un guichet unique qui met à la disposition des acteurs publics et privés les ressources informatiques et les services associés (formation, mentorat, assistance technique, accès à des outils, jeux de données et modèles, services d'incubation et accès à des VCs) dont ils ont besoin pour accélérer leurs travaux de recherche sur le développement et le déploiement de l'IA.

Sites web : <https://www.genci.fr> et <https://aifactory-france.eu>

Contact presse : Nicolas Belot I nicolas.belot@genci.fr I +33 7 60 99 95 10

À propos du CNRS

Acteur majeur de la recherche fondamentale à l'échelle mondiale, le Centre national de la recherche scientifique (CNRS) est le seul organisme français actif dans tous les domaines scientifiques. Sa position singulière de multi-spécialiste lui permet d'associer les différentes disciplines scientifiques pour éclairer et appréhender les défis du monde contemporain, en lien avec les acteurs publics et socio-économiques. Ensemble, les sciences se mettent au service d'un progrès durable qui bénéficie à toute la société.

Site web : www.cnrs.fr

Contact presse : Manon Landurant I manon.landurant@cnrs.fr I + 33 6 13 84 40 63

À propos de Scaleway

Scaleway est le fournisseur européen de cloud et d'IA souverains, proposant une plateforme sécurisée, transparente et durable. Nous accompagnons les organisations de toutes tailles grâce à des technologies ouvertes et indépendantes et une innovation continue, afin qu'elles puissent concevoir et déployer leurs projets à leur propre rythme et selon leurs priorités.

Filiale du Groupe Iliad, Scaleway allie des décennies d'expertise dans les infrastructures à l'agilité d'une entreprise technologique de nouvelle génération. En s'appuyant sur un réseau de datacenters en constante expansion à travers l'Europe, Scaleway propose un portefeuille complet de services cloud haute performance, allant des machines virtuelles et solutions avancées de gestion des données aux infrastructures cloud-native et aux supercalculateurs optimisés pour l'IA. En défendant les standards ouverts et en opérant dans un cadre pleinement européen, Scaleway offre un environnement cloud sécurisé et transparent, capable de répondre aux exigences les plus élevées en matière de souveraineté numérique.

Website : <https://www.scaleway.com/fr/>

Contact presse : presse@scaleway.com

À propos de la RATP



3^e opérateur de transports urbains avec plus de 14 millions de voyages assurés chaque jour dans le monde, le groupe RATP est un acteur majeur du secteur, présent dans 16 pays sur 5 continents. Il exploite au quotidien 9 modes de transport (métro, bus urbain et interurbain, tramway, ferroviaire, sightseeing, câble, navette maritime, transport à la demande et véhicule autonome). Il est aussi présent, en partenariat, sur 3 autres modes (covoiturage, autopartage, vélo en libre-service). Expert reconnu dans les domaines de la gestion d'infrastructures et de l'ingénierie, le Groupe propose également, à travers ses filiales spécialisées, une large palette d'expertises connexes, allant des services urbains (immobilier, télécom, logistique, nouvelles mobilités, énergie) à la sûreté, en passant par les solutions d'information voyageur et de billettique ou le déploiement de surfaces commerciales dans les espaces de transport. En Île-de-France, la RATP exploite depuis 75 ans l'un des réseaux multimodaux les plus denses au monde. Grâce à plus de 73 500 collaborateurs, dont près d'un quart à l'international, le groupe RATP conçoit, met en œuvre et fait vivre au quotidien des solutions de mobilité et des services innovants pour une ville durable et plus humaine. @RATPGroup

Contact presse : servicedepresse@ratp.fr I + 33 1 58 78 37 37