

Ambassade de France en Inde  
Service économique de New Delhi

Affaire suivie par Feli Visco, Marie Khater  
Avec les compléments de l'IFI et de M. Francis Rousseau (ETI iSPIRT)  
Visa : Benoît Gauthier  
24-168

## INDE

### La place de l'Inde dans la chaîne de valeur de l'Intelligence Artificielle

*En s'appuyant sur les forces de son écosystème numérique (taille du marché, réservoir de talents, succès de ses infrastructures numériques publiques), l'Inde veut se positionner comme développeur sur l'intelligence artificielle (IA) mais également comme consommateur d'applications pour soutenir son développement socio-économique, en particulier dans certains secteurs clés comme l'agriculture, la santé ou l'éducation. De nombreux défis restent à relever sur la gouvernance, le développement de puissance de calcul ou sur les recompositions à venir du marché du travail.*

#### 1. Le gouvernement a adopté récemment une stratégie sur l'IA (IndiaAI Mission), dotée d'un budget d'1,2 Md USD

Si les autorités indiennes travaillent depuis 2017 à l'élaboration d'une stratégie nationale, le programme **IndiaAI Mission a finalement été approuvé en mars 2024 et est doté d'un budget d'1,2 Md USD sur cinq ans (103 Md INR)**. Pilotée par le ministère de l'électronique et des technologies de l'information (MeitY), cette stratégie s'appuie sur un rapport d'experts publié à l'automne 2023<sup>1</sup>.

L'IndiaAI Mission repose sur sept piliers :

- 1) **Infrastructures et capacités de calcul** : 50 % du budget de la mission (547 MUSD) seront dédiés à la construction de ce socle indispensable ; le gouvernement souhaite acquérir une quantité très importante de cartes graphiques permettant de paralléliser les calculs (10 000 GPU)<sup>2</sup> via un partenariat public-privé. Ces cartes graphiques devraient permettre d'approvisionner le premier niveau de l'infrastructure de calcul à trois niveaux, dédiée à l'IA (*voir annexe I*).
- 2) Un **centre d'innovation** pour le développement et le déploiement de grands modèles multimodaux de traduction (*large multimodal models*) en s'appuyant sur des données textuelles mais également audio et vidéos et de modèles de fondation d'IA dans des secteurs critiques. Le développement d'applications de modèles de traduction automatique (programme national de traduction en langues vernaculaires - Bhashini) dans les 22 langues officielles de l'Inde est actuellement en phase de mise en œuvre et de déploiement (*voir annexe II*).
- 3) Une **plateforme publique permettant l'accès à des bases de données non personnelles (India Datasets platforms)**. Sur le modèle de l'India Stack (ensemble d'infrastructures numériques publiques)<sup>3</sup>, l'Inde mise sur un programme unifié d'échange de données pour gérer l'accès, les licences et la standardisation des données, des métadonnées et des API hébergées par divers ministères. Des accords ont déjà été conclus avec d'autres ministères pour que leurs données soient placées sur cette plateforme et que les chercheurs et les jeunes entreprises indiennes puissent y avoir accès moyennant paiement. **Les modalités de gestion de ces données sur la plateforme publique sont en cours de définition**. Un projet de cadre national de gouvernance des données a été publié en mai 2022, décrivant les principes et les lignes directrices en la matière. Un bureau national de gestion des données sera créé qui aura pour mission de réguler l'usage des données anonymisées non personnelles et d'en contrôler l'accès. D'autres initiatives, concurrentes par certains aspects, se développent en parallèle notamment le programme National Data and Analytics Platform du think tank gouvernemental NITI Aayog qui regroupe près de 85 bases de données issues de l'administration centrale et des États fédérés ou l'initiative DEPA training du think tank iSPIRT (cf. infra).
- 4) **Applications pour un développement économique inclusif** : Le MeitY cherche à développer 25 applications d'IA dans différents secteurs. Il s'agit soit d'agrèger des solutions existantes développées

par différents ministères et États fédérés, soit d'identifier des problèmes auxquels l'IA peut apporter des solutions.

- 5) **Formation** : L'objectif est de renforcer l'offre de formation en IA sur l'ensemble des niveaux d'éducation supérieure et pour la formation professionnelle. En particulier, une initiative vise à former des travailleurs pour les emplois subalternes tels que l'annotation de données dans les villes de niveau 2 et 3. Sous la direction du ministère indien de l'enseignement supérieur, trois centres d'excellence de l'IA dans les domaines de l'agriculture, de la santé et des villes durables seront créés et disposeront d'un budget de près de 12 M USD (990 M INR) sur cinq ans, (*voir annexe III*).
- 6) **Financement des startups deeptech** : Le MeitY soutiendra les start-ups deeptech par des programmes de financements publics déjà existants et leur connexion avec un réseau national d'incubateurs.
- 7) **La gouvernance de l'IA** est le dernier pilier de l'IndiaAI Mission. Le gouvernement veut promouvoir une IA sûre et de confiance, en utilisant les applications développées par les pouvoirs publics comme des modèles de bonne conduite et en développant des outils techniques permettant de garantir un certain cadre éthique aux applications de l'IA.

Un appel à manifestation d'intérêt a été lancé pour créer des outils techniques et des cadres d'évaluation afin de promouvoir un développement et un déploiement responsables de l'IA sur des thèmes variés (désapprentissage machine, génération de données synthétiques, outils d'équité algorithmique, stratégies d'atténuation des biais de l'IA, cadres éthiques de l'IA, stratégies d'amélioration de la protection de la vie privée, cadres explicatifs de l'IA, certification éthique de l'IA, cadre de test de la gouvernance de l'IA, outils d'audit algorithmique).

L'Inde poursuit ainsi son approche de la gouvernance basée sur des solutions techniques (par exemple de « privacy by design »).

## **2. L'Inde est bien positionnée sur les modèles spécialisés et sectoriels, qui s'appuient pour certains sur le succès de ses infrastructures numériques publiques (India Stack)**

**L'Inde a développé un ensemble d'infrastructure numérique publique (identité, paiement et partage de données basé sur le consentement) – India Stack<sup>3</sup> -, dont la diffusion a été un réel succès en matière d'inclusion économique et sociale par le numérique. L'Inde souhaite capitaliser sur ce succès, et envisage les applications en IA comme l'étape supérieure à ces infrastructures.**

Le think tank iSPIRT, à l'origine du troisième pilier de l'India Stack, le cadre de partage de données en temps réel basé sur le consentement (DEPA – Data Empowerment and Protection Architecture), développe l'utilisation de ce cadre pour l'échange de données d'apprentissage pour l'IA – *DEPA training*.

*DEPA training* repose sur trois concepts clés : (i) un contrat entre le fournisseur des données et l'utilisateur de ces données pour entraîner un modèle, (ii) la mise en place de *clean room* qui assure l'utilisation des données sur des infrastructures protégées et sécurisées et (iii) la confidentialité différenciée qui est une approche en traitement des données permettant leur anonymisation.

La France a détaché depuis octobre 2023 un expert technique international au sein d'iSPIRT qui contribue à cette approche jugée prometteuse d'infrastructure numérique publique au service du bien commun.

La stratégie de l'Inde qui se focalise à la fois sur la création de bases de données et d'applications concrètes dans des secteurs à fort enjeu en Inde (agriculture, santé, ville durable), en s'appuyant sur la taille du marché indien et l'appétence pour le numérique de la population indienne, est largement plébiscitée dans la sphère politique et économique, notamment M. Nandan Nilekani, co-fondateur du groupe informatique indien Infosys, à l'initiative du système d'identité numérique Aadhar<sup>4</sup>. Selon M. Nilekani, l'Inde pourrait devenir le centre mondial des cas concrets d'utilisation de l'IA pour servir le développement économique et social.

**Le développement actuel du marché est en ligne avec ce discours politique et stratégique, orienté sur les applications concrètes de l'IA.** Certes, l'Inde développe des cas d'usage de modèles de fondation (LLM) qui prennent en compte la diversité linguistique du pays (programme de traduction en langues vernaculaires - Bhashini - *voir annexe II*) mais ces LLM reposent en général sur les modèles de fondation des Big Tech américains. La diffusion de l'IA est aujourd'hui essentiellement portée par trois types d'acteurs en Inde :

*les Global Capacity Centres* des grands groupes qui souhaitent se positionner comme centres d'excellence en IA appliquée, un écosystème de start-up dynamique essentiellement orienté sur les applications en particulier pour les entreprises et un écosystème très dynamique de non-profit Tech comme Whadhwani AI<sup>5</sup>. Les centres de compétences mondiaux (*Global Capability Centres*) sont des unités

installées dans un pays donné par une multinationale pour remplir une série de fonctions stratégiques servant les activités mondiales du groupe. L'Inde est une destination privilégiée de ces centres, avec plus de 1 500 entreprises internationales en ayant installé un dans le pays, représentant près d'1,7 million d'emplois sur l'année fiscale 2023<sup>6</sup>. Depuis plusieurs années, ces centres développent des applications d'IA pour améliorer l'efficacité opérationnelle des grands groupes<sup>7</sup>. C'est notamment le cas pour les entreprises françaises, comme Airbus, Capgemini, Michelin ou encore Schneider Electric<sup>8</sup>.

Airbus dispose à Bangalore de son deuxième plus grand centre mondial de gestion des systèmes d'information avec près de 1 200 salariés, avec l'IA et le Big Data au cœur de ses activités.

Michelin dispose d'un centre sur les technologies de l'information à Pune depuis 2014, spécialisé dans les technologies émergentes dont l'analyse de données, l'IA et l'IoT.

Capgemini dispose d'un centre d'excellence spécialisé en intelligence artificielle qui développe des solutions dans le domaine de la reconnaissance visuelle, du traitement automatisé du langage naturel, de l'intelligence artificielle conversationnelle et de l'automatisation. Schneider Electric dispose de trois centres mondiaux spécialisés en IA - en Inde, en France et aux États-Unis. En miroir, le groupe informatique indien TCS (Tata Consulting Services) a annoncé lors de Choose France 2024 la création d'un centre d'excellence en IA à Paris<sup>9</sup>. Selon l'économiste S Anklesaria Aiyar, l'IA est une opportunité pour une montée en gamme des activités développées dans ces GCC.

L'écosystème de start-up indiennes sur l'IA est également très dynamique. Selon un rapport d'EY<sup>7</sup>, 66% des cinquante premières start-ups indiennes en termes de valorisation utilisent déjà l'IA et le boom de l'IA générative dans cet écosystème s'oriente vers quatre principales thématiques d'applications pour les entreprises : (i) marketing et ventes, (ii) support clientèle, (iii) création de contenus pour les médias, (iv) applications sectorielles spécifiques (découverte de médicaments en santé par exemple).

### **Ce développement axé cas d'usage est cohérent avec les caractéristiques de l'économie numérique indienne.**

D'une part, l'Inde dispose d'un **marché numérique profond pour la diffusion de ces modèles** – 700 millions d'utilisateurs d'internet soit le deuxième plus grand marché au monde. D'autre part, l'Inde dispose **d'un important réservoir de talents** avec 600 000 personnes formées dans ce domaine, faisant du pays le deuxième pôle mondial dans ce secteur.

**En revanche, le pays n'est pas très bien placé en matière de R&D.** Selon une étude de Change Engine, l'Inde est au quatorzième rang en matière de contribution aux articles de recherche des dix plus grandes conférences, classement dominé par les États-Unis et la Chine (30 % et 23 % respectivement)<sup>10</sup>, tandis que la France est au onzième rang. Le centre "Human-Centered Artificial Intelligence" de l'université de Stanford considère qu'aucun modèle majeur d'IA n'a été développé en Inde contre 61 aux États-Unis, 15 en Chine et 8 en France, les trois pays en tête du classement<sup>11</sup>. La recherche en IA en Inde souffre d'un manque de moyens et d'incitations à la qualité de la recherche. L'Inde est mieux placée en revanche dans la recherche appliquée, dont les bénéfices à long terme sont moindres mais qui contribuent à la diffusion de l'IA dans l'économie.

L'écosystème de recherche indien se structure progressivement avec des départements dédiés dans les IIT (Indian Institute of Technology) de Bombay, Madras, Delhi, Hyderabad et Kharagpur, les universités scientifiques mais aussi les IIITs (Indian Institutes of Information Technology) dont c'est la spécialité, comme celui de Bangalore. La coopération bilatérale entre l'Inde et la France en mathématiques appliquées est riche et s'ouvre progressivement à l'IA et à ses applications sectorielles (santé<sup>12</sup>, agriculture). Le Centre indo-français de mathématiques appliquées (IFCAM), laboratoire international de recherche du CNRS, a été créé en 2012 à Bangalore dans le but d'améliorer et de développer les collaborations en mathématiques appliquées (68 projets, neuf écoles d'été ou d'hiver, sept ateliers, quatre-vingt-dix facultés, quarante doctorants, dix post-doctorants de 42 institutions différentes des deux pays ont participé à l'IFCAM). Le CEFIPRA (Indo-French Centre for Promotion of Advanced Research) soutient actuellement 26 projets de recherche dans le domaine des sciences informatiques fondamentales et appliquées et des mathématiques, avec l'émergence progressive de projets transdisciplinaires impliquant l'IA (santé, agriculture). Le CEFIPRA soutient également un programme de coopération entre l'INRIA et le DST en sciences et technologies du numérique, lancé en 2013, qui met l'accent sur des domaines tels que le Big Data, les systèmes cyber-physiques, le calcul haute performance, les systèmes embarqués, le calcul fiable et évolutif, l'informatique pour la biologie et les sciences de la vie, la cybersécurité et l'apprentissage automatique. Enfin, l'Inria, l'IIT Delhi et Naval Group portent un programme de coopération sur les applications potentielles de l'IA en santé, cybersécurité, robotique, systèmes d'information et machine learning.

### **Enfin, le pays fait partie des géographies bénéficiant d'investissements privés conséquents en IA.**

En 2023, 1,39 Md USD a été investi par des acteurs privés en Inde dans des activités liées à l'IA, positionnant le pays au 10ème rang mondial très loin derrière les États-Unis (près de 70 Mds USD) mais sur un ordre de grandeur proche de pays comme la France, l'Allemagne ou Israël (entre 1,5 Md USD et 2 Md USD)<sup>11</sup>. Le financement de start-ups en Inde reste très concentré sur le "early funding" et a diminué depuis 2023 en raison d'un mouvement de correction de marchés survalorisés<sup>13</sup>. Cette baisse touche également les start-ups en IA en Inde, contrairement au reste du monde où ce marché a plutôt bien résisté au ralentissement des investissements dans la Tech. Selon l'AIM (Artificial Intelligence Market) Research, les start-ups en IA en Inde ont levé 560 M USD en 2023, un montant deux fois moins élevé qu'en 2022<sup>14</sup> et la tendance se poursuivrait depuis le début de l'année 2024<sup>15</sup>.

### **3. Des défis majeurs seront à relever pour que l'Inde puisse tirer parti du développement mondial de l'IA**

#### **a. La gouvernance de l'IA : un difficile équilibre entre réguler l'utilisation de l'IA en imposant un contrôle de l'État et ne pas brider l'innovation**

**Au niveau international**, notamment au sein du PMIA (Partenariat Mondial sur l'IA) où l'Inde a eu la présidence cette année et lors du UK AI Safety Summit, **l'Inde soutient un développement inclusif et responsable de l'IA, qui respecte des normes éthiques**. Le gouvernement indien insiste en particulier sur la nécessité de développer des modèles qui prennent en compte la diversité culturelle des pays (en évitant les biais liés au recours à des bases de données non représentatives au niveau mondial).

**Sur le plan interne, il n'existe pas aujourd'hui de régulation spécifique sur l'IA**. En 2021, le think tank gouvernemental NITI Aayog, en partenariat avec le *World Economic Forum*, a publié une étude sur l'utilisation responsable des technologies émergentes qui n'a pas eu de suites. NITI Aayog propose une approche basée sur les risques, selon laquelle les applications à faible risque seraient auto-régulées tandis que pour les applications à risque élevé, l'État imposerait des principes éthiques. La méthodologie d'évaluation du risque n'est pas précisée. Ces principes éthiques portent sur sept composantes : la sûreté et la fiabilité, l'inclusion et la non-discrimination, l'égalité, la protection de la vie privée et la sécurité, la transparence, la responsabilité et la protection et le renforcement des valeurs humaines positives.

**L'espace numérique est aujourd'hui régulée par l'IT Act de 2000, modifié en 2008, et les différents décrets d'application (IT Rules)**. Le MeitY a annoncé en 2022 vouloir mettre à jour le cadre réglementaire avec un *Digital India Act* visant à réguler les nouvelles technologies dont l'IA mais au vu des difficultés rencontrées, **le gouvernement indien s'orienterait vers des modifications à la marge du cadre réglementaire existant**. Deux principaux amendements aux *IT Rules* ont été notifiés récemment et portent essentiellement sur les contenus publiés dans l'espace numérique. Par ailleurs, le Parlement indien a adopté en août 2023 une première loi sur la protection des données à caractère personnel, qui doit encore entrer en vigueur et dont de nombreux aspects doivent être précisés dans des règlements à venir. Enfin, l'Inde a publié en mars 2024, une première version d'une loi sur la concurrence dans l'économie numérique (*Digital Competition Bill*)<sup>16</sup> qui vise à mettre en place une réglementation sélective *ex ante* des grandes entreprises numériques qui offrent des services numériques de base (services d'intermédiation et d'agrégation de contenus) et qui est pour le moment soumise à la consultation publique.

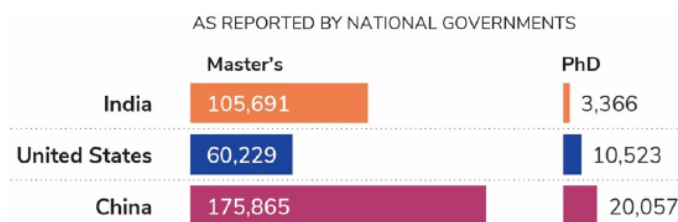
**La faiblesse du cadre réglementaire sur la protection des données personnelles est un des facteurs qui alimentent la prégnance des risques cyber dans le pays**. L'Inde a été le 5<sup>ème</sup> pays le plus touché par des violations de données en 2023 (derrière les États-Unis et la France notamment), tandis que ces violations<sup>17</sup> auraient coûté en moyenne 2,4 M USD aux entreprises indiennes en 2023 contre 9,4 M USD aux États-Unis et 4,1 M USD en France, selon IBM. Par ailleurs, l'Inde présente le taux de fraude le plus élevé au monde sur les paiements numériques instantanés<sup>18</sup>.

#### **b. L'impact incertain de l'IA sur le marché de l'emploi indien, avec un risque de creusement des inégalités**

**L'Inde dispose d'un réservoir main d'oeuvre formé en IA parmi les plus importants au monde**. Selon une estimation réalisée par le cabinet de conseil BCG et la première fédération d'entreprises indiennes du secteur IT, 600 000 personnes seraient formées en Inde chaque année en IA, performance plaçant le pays au deuxième rang mondial dans ce secteur<sup>6</sup>. **Il convient toutefois de nuancer ce classement par le niveau d'expertise de ce réservoir de main d'oeuvre**. En effet, si on considère comme proxy du niveau

d'expertise en IA le niveau de diplôme, l'Inde est moins bien positionnée au niveau mondial. Si le pays dispose d'un nombre d'ingénieurs diplômés (niveau master) dans des domaines relatifs à l'IA largement supérieur aux Etats-Unis, le pays est loin derrière la Chine et les États-Unis pour les doctorats (cf. graphique ci-dessous). Ce constat se retrouve également sur la recherche en IA (cf. supra).

Table 1: Engineering graduates, 2016–17.



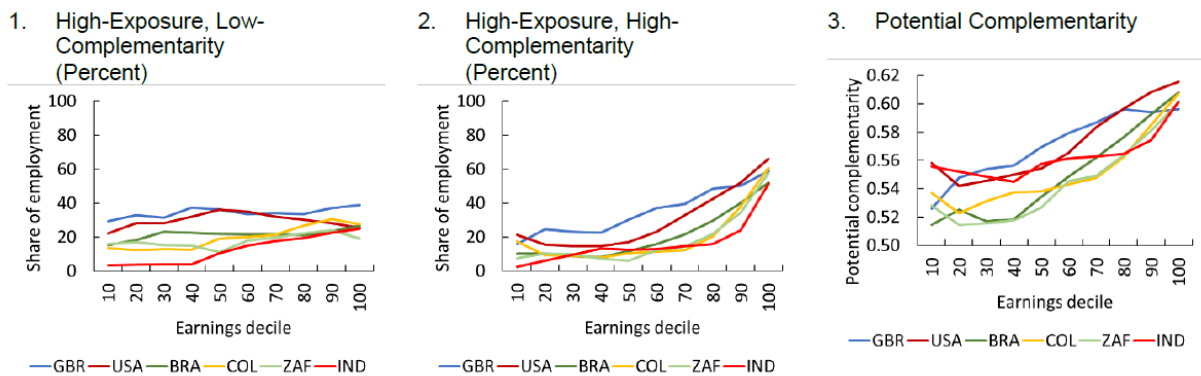
Source: All India Survey on Higher Education, Ministry of Education in China, and the National Center for Education Statistics in the United States.

Source: "Mapping India's AI Potential", CSET (Center for Security and Emerging Technology) Data Brief, March 2021

Les services IT représentent un secteur d'activité majeur en Inde, avec des centres de multinationales initialement concentrés sur des activités à faible valeur ajoutée qui sont progressivement montés en gamme. Une partie de ces activités à faible valeur ajoutée comme les centres d'appels ou le développement informatique pourrait être automatisée grâce à l'IA, engendrant des destructions d'emplois. **Ainsi, le développement de l'IA et en particulier des applications de l'IA générative pourrait entraîner une forte recomposition de l'emploi au sein de l'industrie IT, nécessitant des politiques très volontaristes de formation pour permettre aux salariés de monter en compétences.** Par exemple, M. K. Krithivasan, directeur général de Tata Consulting Services, un des plus grands groupes indiens de services informatiques, alerte sur les risques de destructions d'emplois liés au développement de l'IA pour les centres d'appel en Inde<sup>19</sup>.

**Au-delà de la recomposition du marché de l'emploi dans les services informatiques, qui ne pourra être réussie qu'avec l'appui de politiques de formation, le développement de l'IA pourrait renforcer les inégalités sur le marché du travail de manière prononcée en Inde selon une étude récente du FMI<sup>20</sup>.** Du fait d'une forte concentration de la main d'oeuvre indienne dans le secteur agricole, l'Inde présente la proportion la plus faible d'emplois exposés à l'IA définis comme les emplois dont les occupants peuvent mobiliser l'IA pour réaliser leur activité. Parmi les pays émergents, l'Inde obtient le taux le plus faible à 26 % quand le Brésil atteint 41 % et les économies avancées 60 %. Par ailleurs, l'IA pourrait aggraver les inégalités sur le marché du travail, puisque des synergies ne semblent apparaître que pour les métiers déjà les plus rémunérateurs, et ce d'une manière bien plus notable qu'ailleurs dans le monde (cf. figure ci-dessous).

**Figure 4. Share of Employment in High-Exposure Occupations and Potential Complementarity by Income Deciles**



Sources: American Community Survey; Gran Encuesta Integrada de Hogares; India Periodic Labour Force Survey; Labour Market Dynamics in South Africa; Pesquisa Nacional por Amostra de Domicilios Continua; Pizzinelli and others (2023); UK Labour Force Survey; and IMF staff calculations.

Note: Panel 1 shows the employment share in jobs with high exposure but low complementarity, and panel 2 presents the employment share in jobs with high exposure and high complementarity, each categorized by income deciles. Panel 3 shows the potential AI occupational complementarity from Pizzinelli and others (2023), averaged and grouped by income deciles. Country labels use International Organization for Standardization (ISO) country codes.

Source: Cazzaniga and others. 2024. « Gen-AI: Artificial Intelligence and the Future of Work.» IMF Staff Discussion Note SDN2024/001, International Monetary Fund, Washington, DC.

**c. Les infrastructures de calcul : de grandes ambitions indiennes mais un retard technologique net sur les puces et des besoins massifs de développement des centres de données (data centers)**

**L'Inde ne dispose pas des capacités de production de puces essentielles pour le développement de l'IA, en particulier des processeurs graphiques (GPU), ni d'entreprise spécialisée dans ce domaine.**

Une industrie de semi-conducteurs et puces commence à se développer sous l'impulsion volontariste des pouvoirs publics, mais à ce stade essentiellement dans l'assemblage de puces moins avancées, utilisées dans certaines industries (automobile, téléphonie)<sup>21</sup>.

**Le pays se concentre donc sur la sécurisation de son approvisionnement en GPU pour répondre aux besoins du marché domestique, avec des annonces de partenariat à la fois du côté de l'Etat indien (cf. supra avec l'annonce de l'achat à venir de 10 000 GPU) et des grands groupes (Reliance, Tata et le fournisseur de capacités de calcul Yotta)<sup>22</sup>.**

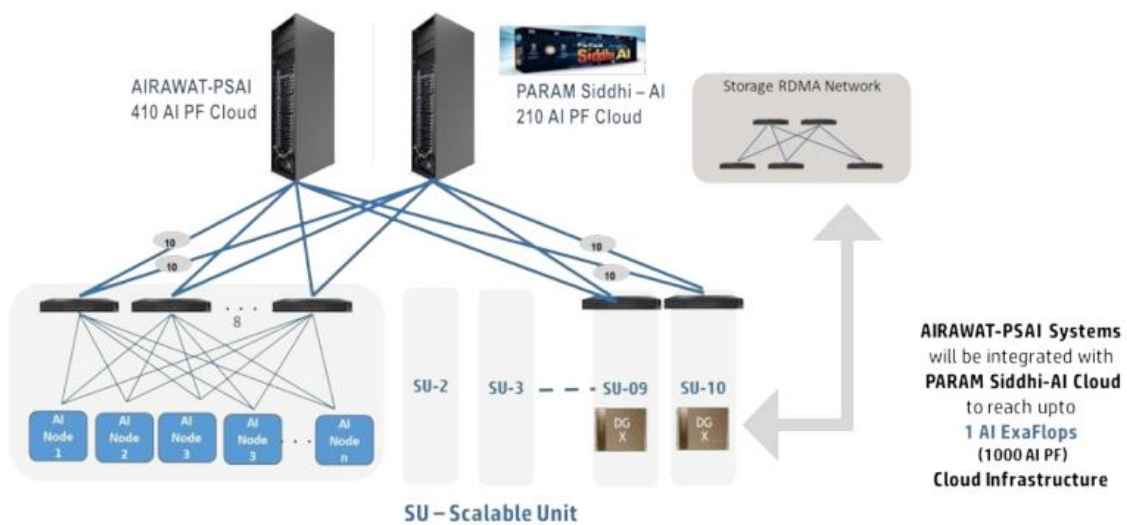
Ainsi, le développement de capacités de calcul se structure, à la fois du côté des infrastructures publiques avec une stratégie désormais bien définie (*voir annexe II*) et du côté du secteur privé avec des grands groupes qui achètent directement des GPU pour leurs besoins internes.

Pour faire face à la croissance de la demande de capacités de calcul et de stockage de données, l'Inde prévoit de développer les data center. Un récent rapport de l'entreprise américaine dans les services immobiliers commerciaux Cushman & Wakefield pointe des besoins additionnels par rapport aux projets déjà engagés. Au-delà de 2,32 GW de data center déjà prévus d'ici 2028, 1,7 à 3,8 GW supplémentaires seraient nécessaires, ces chiffres étant à mettre en perspective avec les 977 MW actuellement installés<sup>23</sup>. Les data center devront être mieux répartis sur le territoire, Mumbai, Chennai, Delhi et Hyderabad concentrant aujourd'hui 90% de ces infrastructures de calcul de de stockage de données. Ce marché pourrait bénéficier à des entreprises françaises, dont OVHCloud qui a inauguré en 2023 son premier data center en Inde à Bombay.

**En 2015, l'Inde a lancé une mission nationale dédiée aux supercalculateurs** (National Supercomputing Mission), doté d'un budget de 489 M EUR afin de développer un réseau de 24 supercalculateurs dans le pays sur une période de 7 ans. Eviden (anciennement Atos) a remporté un contrat pour développer des supercalculateurs dans le pays et a fourni 14 des 24 ordinateurs. L'un d'entre eux, le PARAM Siddhi-AI, figure actuellement au 185<sup>ème</sup> rang du Top500 des supercalculateurs mondiaux. PARAM Siddhi AI est un supercalculateur de 210 pétaflops avec 42 nœuds, des processeurs NVIDIA DGX A100, AMD EPYC 7742 64C 2,25 GHz, NVIDIA A100 et des interconnexions Mellanox HDR Infiniband.

**Le supercalculateur le plus puissant en Inde, qui figure à la 110<sup>ème</sup> place mondiale**, a été construit par l'entreprise indienne Netweb Technologie et a été inauguré en mai 2023 (Airawat PSAI). Airawat-PSAI est un supercalculateur de 200 pétaflops de précision mixte avec 82 nœuds, NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHZ, processeurs NVIDIA A100 et une interconnexion INFINIBAND HDR. Il utilise le logiciel Ubuntu 20.04.2 LTS et consomme 25 000 kW d'énergie.

**Ensemble, ces deux systèmes fournissent une puissance de calcul de 410 pétaflops d'IA et le gouvernement indien prévoit d'augmenter cette puissance pour atteindre une puissance de calcul exaflopique.** L'infrastructure d'IA est décrite dans la figure ci-dessous :



Au-delà de l'atteinte d'une puissance exaflopique, l'Inde souhaite développer une infrastructure de calcul et d'inférence pour l'IA à trois niveaux. Cette architecture a été présentée dans le rapport d'experts publiés en octobre 2023 dans le cadre de l'initiative FutureLab AI qui s'appuiera sur des partenariats public-privé :

1. Niveau 1 - Installation d'infrastructure de calcul haut de gamme avec 40 Exaflops de puissance de calcul pour l'IA, un cluster de 10 000 GPU et 200 pétaoctets (200 x 10<sup>15</sup> octets) de stockage. Ce niveau sera utilisé pour l'entraînement de modèles nécessitant des capacités de calcul massives.
2. Niveau 2 – Capacités de calcul de gamme moyenne, réparti sur quatre sites géographiques, avec une puissance de calcul d'IA de 3 exaflops, 7 000 GPU et 400 pétaoctets de stockage. Ces centres de niveau 2 seront utilisés pour l'entraînement mais également pour l'inférence en s'appuyant sur leur répartition géographique. Ce niveau 2 bien réparti géographiquement permettra de réduire les temps de latence et d'améliorer l'accessibilité aux capacités de calcul et d'inférence sur l'IA sur tout le territoire indien.
3. L'installation d'infrastructures de calcul en périphérie de niveau 3 réparties sur 12 sites avec une puissance de calcul pour l'IA de 500 pétaflops, 7 500 GPU chacun et 240 pétaoctets de stockage. Ces centres régionaux seront utilisés pour l'entraînement et l'inférence des modèles à l'échelle plus locale.



## Annexe II – Le programme Bhashini, un modèle de traduction automatique dans les différentes langues indiennes

[Bhashini](#) est l'un des projets phares du MeitY lancé en 2022 et vise à mettre à la disposition des start-ups et des acteurs du secteur privé des ensembles de données audio et textuelles provenant des 22 langues nationales de l'Inde, afin qu'ils puissent développer des applications innovantes. The [AI4Bharat](#), laboratoire de recherche de l'IIT Madras, l'une des plus grandes écoles d'ingénieurs en Inde, développe des outils, des modèles et des applications en open source et met en place également une base de données pour les langues indiennes dans le cadre du projet Bhashini. Le centre de recherche a développé quatre outils open-source pour l'utilisation de plusieurs langues indiennes, comme suit :

1. Shoonya, une plateforme open source pour l'annotation et l'étiquetage des données.
2. Chitralekha pour le sous-titrage de vidéos.
3. Karvy pour informer les utilisateurs sur les enjeux de protection de la vie privée et des données.
4. Anuvaad pour la traduction de documents judiciaires.

A terme, cet outil permettra de fournir des services publics ciblés aux citoyens indiens dans leur propre langue. Par exemple, l'État a collaboré avec Microsoft pour développer une plateforme de solution d'IA conversationnelle [Jugalbandi](#) qui utilise Bhashini et ChatGPT, permettant des échanges dans dix des 22 langues nationales et fournissant des informations sur 171 programmes gouvernementaux.

Sur la construction de bases de données dans les différentes langues indiennes, le MeitY a déjà collecté de données par crowdsourcing à partir de citoyens qui fournissent des traductions orales et écrites et de personnes déployées sur le terrain qui recueillent des enregistrements vocaux. Actuellement, 100 000 heures de données en langue indienne ont été collectées et le MeitY souhaite aller plus loin en collectant des données vidéos. Un protocole d'accord avec le ministère de la radiodiffusion a été signé pour avoir accès à d'anciennes vidéo diffusées sur la chaîne nationale indienne Doordarsha. Au total, le programme Bhashini a permis déjà permis de développer 300 à 400 modèles en langue indienne.

La diversité culturelle et linguistique de l'Inde appuyée par l'effort de l'Etat indien pour créer des bases de données riches en open source est source d'opportunités pour l'écosystème de start-ups indiennes, très dynamique dans ce domaine. Par exemple, la start-up indienne Ola (concurrente d'Uber) a lancé Krutrim, première licorne indienne dans le domaine de l'IA<sup>24</sup>, qui propose un LLM capable de générer du texte en 10 langues indiennes. Si des rumeurs ont circulé sur le fait que Krutrim propose un LLM adapté à partir des modèles proposés par OpenAI, aucune information précise sur les données utilisées, les méthodes d'apprentissage du modèle, les capacités de calcul utilisées a été communiquée officiellement. D'autres start-ups proposent des modèles génératifs adaptés aux langues vernaculaires indiennes dont la start-up Sarvam AI qui a publié le modèle OpenHathi, le premier LLM en langue hindi adaptée à partir de l'architecture des modèles de Meta. Le modèle OpenHathi a été ajusté pour utiliser des données audios, dont les applications de chatbot sont très utiles dans un pays comme l'Inde où l'alphabétisation n'est pas



encore universelle (taux d'alphabétisation de près de 78% et 70% pour les femmes contre 87 % au niveau mondial).

Ces solutions linguistiques en cours d'élaboration ne sont toutefois pas développées de manière totalement souveraine puisqu'elles s'appuient sur des solutions cloud et des programmes en IA fournis par des Big Tech telles que Microsoft et Amazon. Le centre de recherche AI4Bharat collabore avec Microsoft Research et utilise le cloud et les outils d'IA d'Azure, tandis que la solution Jugalbandi est le fruit d'une collaboration entre le gouvernement indien et Microsoft. La plupart des modèles développés par les start-ups indiennes s'appuient sur les modèles LLaMA de Meta et GPT de Microsoft.

### Annexe III – Présentation des centres d'excellence sur l'IA ciblés sur des secteurs (santé, agriculture et ville durable)

Sous la direction du ministère indien de l'enseignement supérieur 3 centres d'excellence de l'IA dans les domaines de l'agriculture, de la santé et des villes durables seront créés et disposeront d'un budget de près de 12 M USD (990 M INR) sur 5 ans. 11 cohortes ont déjà été sélectionnées lors de la phase 1 pour passer à l'étape de la preuve du concept, dont 4 dans le domaine de la santé, 5 dans le domaine des villes durables et 2 dans le domaine de l'agriculture. Chaque cohorte recevra une subvention de 20 M INR ou 239 000 USD.

The infographic is titled "AI Centres of Excellence (AI-CoE)" and is an initiative of the Ministry of Education. It features the logos of the Ministry of Education and the Indian Institute of Technology Jammu. The main text states: "In a leap towards leveraging AI for the public good, MoE launched an ambitious program to establish AI Centre of Excellence (CoE) in Agriculture, Health, and Sustainable Cities. 11 cohorts have progressed in Phase 1 to showcase their Proof-of-Concepts (PoC). PoC phase will culminate in establishment of three CoEs aligned with Government of India vision to 'Make AI in India and Make AI work for India.'" The infographic is divided into three main sections: "Sustainable Cities", "Health", and "Agriculture".

**AI Centres of Excellence (AI-CoE)**  
An Initiative of Ministry of Education

**Sustainable Cities**

- Forecasting and Modeling for Urban Sustainability
- Sustainable Infrastructure and Resource Planning, Analysis and Monitoring
- Sustainable Applications for Mobility, Urban Development and Resilience Using AI SAMUDRA
- Generating Casual AI Powered Digital Twin With Reinforcement Learning for Sustainable Urban Transformation
- AI Powered Sustainable Cities

**Health**

- Point of Care Screening Tool for Oral Cancer
- Strengthening Healthcare & Improving Accessibility
- Early Detection and Management of Cardiovascular Diseases & Streamlining Healthcare Facility and Clinical Care
- Coronary Artery Diseases Diagnostic and Prognostic System & AI Solutions for Affordable and Accessible Healthcare

**Agriculture**

- Ensuring Food Security Through AI Assisted Robust Agriculture at Scale
- Leveraging the Potential of AI for Sustainable Agriculture and Food Security: Digital Intelligence for Optimum Use of Water, Nutrients and Pesticides Ensuring National Food Security and Environmental Sustainability

<sup>1</sup> « India AI 2023 », First edition by expert group to MeitY, Octobre 2023. [Lien](#).

<sup>2</sup> Graphic processing units. Pour donner un ordre de grandeur, le supercalculateur le plus performant au monde en 2024, Frontier supercomputer aux Etats-Unis, dispose de 37 000 GPU. [Source](#)

<sup>3</sup> Cf. Note du SER « 23-166 [PE] L'India Stack un modèle de réussite d'infrastructure numérique publique ? ».

<sup>4</sup> Intervention de M. Nandan Nilekani au Global Technology Summit 2023, [lien](#). Selon M. Nilekani, l'IA pourrait permettre d'accélérer le développement économique et social du pays en passant de digital public infrastructure à une approche de digital public intelligence. Les cas d'application de l'IA doivent guider son développement en Inde. M. Nilekani ne considère pas que l'absence de modèles de fondation indiens soit problématique pour assurer la souveraineté numérique du pays. En effet, la course actuelle sur la performance des modèles de fondation va aboutir à faire de ces modèles des commodités.

<sup>5</sup> Wadhvani AI est une organisation à but non lucratif d'IA appliquée, financée par deux milliardaires américano-indiens de la Tech, Suni et Romesh Wadhvani. Wadhvani AI a été lancée en 2018 par le Premier ministre Modi et se concentre sur la création de solutions d'IA appliquée pour l'accès à des biens publics (santé avec notamment des applications pour l'éradication de la tuberculose, éducation, agriculture...). Ces applications concrètes s'appuient généralement sur des modèles open source déjà existants.

<sup>6</sup> Cf. Chiffres issus d'une étude du NASSCOM présentés dans un rapport de KPMG. « GCCs in India : Building resilience for sustainable growth », KPMG et NASSCOM, mai 2024.

<sup>7</sup> Cf. « Is generative AI beginning to deliver on its promise in India », Aldea of India update, EY, Mai 2024.

<sup>8</sup> Airbus dispose à Bangalore de son deuxième plus grand centre mondial de gestion des systèmes d'information, après Toulouse avec près de 1200 salariés, avec l'IA/Big Data au cœur de leurs activités. Michelin a un centre sur les technologies d'information à Pune depuis 2014 qui s'est spécialisé sur les technologies émergentes dont analyse de données, IA et IoT. Capgemini dispose d'un centre d'excellence spécialisé en intelligence artificielle en Inde, qui développe en particulier des solutions dans le domaine de la reconnaissance visuelle, du traitement automatisé du langage naturel, de l'intelligence artificielle conversationnelle et de l'automatisation. Schneider Electric dispose de 3 centres mondiaux spécialisés en IA, en Inde, en France et aux États-Unis.

<sup>9</sup> Source : [TCS To Set Up First-of-its-Kind Human-Centric AI Center of Excellence in France](#).

<sup>10</sup> "India's State of AI Research : India's Performance in Top Conferences as compared to the World", Parth Mahajan, Jitesh Bhasin, Varun Aggarwal, Change Engine, July 2024.

<sup>11</sup> Cf. "The AI Index 2024 Annual Report", AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, Avril 2024.

<sup>12</sup> Un séminaire coorganisé par la Sorbonne et l'AIIMS Delhi est prévu le 16 septembre 2024 « Indo-French Symposium on Imaging, AI, and Genetics in Neurological Disorders ».

<sup>13</sup> Cf Note du SER « 23-255 [Macr] – Le financement de l'innovation en Inde.docx ».

<sup>14</sup> Cf. « AI Startup Funding & Acquisition Report 2024 – India », AIM Research, Février 2024.

<sup>15</sup> « Total Funding for AI startups saw more than 71% dip in 2023 : Report », Business Today, [lien](#), 4 mai 2024.

<sup>16</sup> En première analyse, la délégation de l'UE en Inde considère que ce projet de loi s'inspire largement du Digital Market Act au niveau européen avec néanmoins quelques différences : (i) exclusion de certains secteurs dont le secteur des télécoms, de l'IA générative et des systèmes virtuels sont exclus, (ii) les critères pour définir les Systematically Significant Digital Entreprises restent flous et (iii) pas de prise en compte de l'importance de rendre les plateformes interopérables.

<sup>17</sup> Cette estimation inclut les attaques de rançongiciels, les vulnérabilités liées à l'utilisation du cloud et les usurpations d'identité.

<sup>18</sup> [BIS Papers n° 106 The design of digital financial infrastructure lessons from India](#), December 2019. The Jan Dhan Yojana programme to broaden access to bank accounts and the penetration of mobile phones are key enablers.

<sup>19</sup> « AI could kill off most call centres, says Tata Consultancy Services head », Financial Times, 25 April 2024.

<sup>20</sup> Cf. Cazzaniga and others. 2024. « Gen-AI: Artificial Intelligence and the Future of Work. » IMF Staff Discussion Note SDN2024/001, International Monetary Fund, Washington, DC.

<sup>21</sup> Cf. Note du SER « 23-235 [PE] L'Inde peut-elle devenir un acteur mondial de l'industrie électronique ».

<sup>22</sup> L'entreprise spécialisée dans les centres de données Yotta a annoncé en début d'année 2024 un investissement de près d'1 Md USD à venir pour acheter 32 000 processeurs graphiques de Nvidia (GPU) tandis que Tata et Reliance ont annoncé leur souhait d'acquérir respectivement 8 000 et 16 000 GPU.

<sup>23</sup> « India needs an extra 1.7-3.6 GW data centre capacity by 2028 due to high demand », Hindustan Times, 26 June 2024, [lien](#).

<sup>24</sup> Krutrim a levé début 2024 50 M USD et aurait atteint depuis une valorisation dépassant le milliard d'USD. [Source](#)