



L'ÉTOILE DE L'AFRIQUE

L'AFRIQUE EST-ELLE EN TRAIN D'ENTRAÎNER
L'INTELLIGENCE ARTIFICIELLE

À RACONTER L'AFRIQUE SANS ELLE ?

Données, corpus et représentations : le sujet des dix prochaines années

TZIPPORAH MAYALA

Présidente-Fondatrice de L'Étoile de l'Afrique

Directrice de l'Observatoire de l'Audiovisuel Africain

TRIBUNE

Avec quelles données les machines apprendront-elles ?

Pendant que nous discutons encore de la manière dont le cinéma représente l'Afrique, les intelligences artificielles apprennent déjà à écrire des scénarios, fabriquer des images, générer des voix et reconstituer des univers entiers.

Mais une intelligence artificielle ne se réveille pas un matin avec une connaissance spontanée de Kinshasa, Cotonou, Libreville ou Abidjan. Elle apprend à partir de ce qui a été produit, numérisé, documenté, indexé et rendu accessible.

Une question devient alors urgente : si nous ne construisons pas nous-mêmes nos corpus, nos données et nos infrastructures de connaissance, avec quelle Afrique les machines apprendront-elles à raconter l'Afrique ?

Hier, la bataille était : qui raconte l'Afrique ? Aujourd'hui : qui possède et distribue ses récits ? Demain : avec quelles données les machines apprendront-elles à la raconter ?

MA THÈSE

Une machine n'invente pas un territoire

Il faut le redire simplement, parce que le vocabulaire ambiant entretient la confusion : un système génératif ne connaît pas le monde. Il a appris des régularités dans ce qui lui a été donné à lire, à voir et à entendre. Sa représentation d'un lieu est la moyenne pondérée de ce qui a été écrit et photographié sur ce lieu, par ceux qui avaient les moyens de l'écrire et de le photographier.

Il s'ensuit une conséquence que nous devrions prendre très au sérieux : dans l'univers d'une machine, ce qui n'est pas numérisé n'existe pas. Nos archives non restaurées, nos films non indexés, nos langues peu représentées, nos métiers non documentés ne sont pas sous-représentés : ils sont absents. Et l'absence, dans un modèle statistique, ne produit pas du silence. Elle produit de la substitution — la machine comble avec ce qu'elle connaît.

LE RISQUE

En trois temps

Je ne crois pas au complot. Je crois à la mécanique, et la mécanique est suffisamment inquiétante.

— Premier temps : la sous-représentation

Les corpus disponibles surreprésentent massivement les productions des marchés les mieux équipés en numérisation. Nos cinématographies y figurent à la marge, souvent à travers un petit nombre d'œuvres devenues canoniques. Une machine entraînée ainsi ne connaît pas nos filmographies : elle en connaît quelques titres, et les prend pour le tout.

— Deuxième temps : l'amplification du cliché

Un système génératif reproduit et accentue les régularités qu'il observe. Si les images disponibles d'un territoire sont majoritairement des images de crise, de faune ou de folklore, il produira des images de crise, de faune et de folklore — non par idéologie, mais par statistique. Le cliché n'est plus alors une paresse d'auteur : il devient une propriété du système, reproduite à l'échelle industrielle.

— Troisième temps : la dépossession silencieuse

Nos archives, nos musiques, nos œuvres et nos voix peuvent alimenter des systèmes sans consentement, sans mention et sans rémunération. Notre patrimoine devient alors une matière première gratuite pour des outils que nous devons ensuite acheter. C'est la forme la plus contemporaine, et la plus discrète, d'un mécanisme que nous connaissons bien.

MISE AU POINT

Ce que je ne dis pas

Ce texte n'est pas un procès de la technologie. Je serais mal placée pour l'instruire : nous construisons nous-mêmes des outils qui en dépendent.

Je ne demande pas de refuser l'intelligence artificielle

Elle entre dans nos métiers, elle y restera, et elle rend des services réels — traduction, sous-titrage, repérage documentaire, préparation de dossiers. Refuser l'outil au nom de la souveraineté reviendrait à choisir un désarmement unilatéral. La question n'est pas si, mais à quelles conditions.

Je ne crois pas que la machine remplace l'auteur

Elle propose ; elle ne décide pas. C'est pourquoi tout ce que nous construisons obéit à un principe unique et non négociable : **l'intelligence artificielle assiste, l'humain décide**. Toute chaîne automatisée comporte un point de validation humaine, et la responsabilité éditoriale reste toujours attribuable à une personne.

Je ne confonds pas l'outil et le corpus

Le débat public se concentre sur les modèles. Le vrai enjeu de souveraineté est en amont : les données d'entraînement, leur origine, leurs licences, leur représentativité. On peut utiliser un modèle conçu ailleurs tout en maîtrisant ce qu'on lui donne à apprendre — et c'est là que se joue notre marge de manœuvre réelle.

CE QU'IL FAUT CONSTRUIRE

Cinq chantiers

Il ne s'agit pas d'écrire une déclaration de principe de plus. Il s'agit de poser cinq chantiers concrets, dont aucun n'est hors de portée.

CHANTIER	CE QU'IL SUPPOSE
Numériser	Restaurer et numériser les fonds audiovisuels, les archives sonores et les catalogues nationaux. Ce qui n'est pas numérisé sortira du champ de la connaissance mondiale.
Documenter	Produire les métadonnées : titres, langues, territoires, ayants droit, dates, genres, mots-clés. Une œuvre non indexée reste introuvable pour une machine comme pour un chercheur.
Licencier	Définir les conditions de réutilisation : ce qui peut entraîner un système, à quelles conditions, avec quelle mention et quelle rémunération. Le silence juridique profite toujours à celui qui aspire les données.
Consentir	Recueillir le consentement des auteurs, des interprètes et des personnes filmées pour tout usage d'entraînement. Nous nous l'appliquons à nous-mêmes : les contenus de nos utilisateurs ne servent à aucun entraînement sans accord exprès.
Héberger	Choisir où résident nos corpus, sous quel droit et sous quel contrôle. La souveraineté narrative sera aussi une question d'infrastructure technique.

CE QUE NOUS FAISONS

Prendre notre part

Documenter	L'Observatoire de l'Audiovisuel Africain constitue une base traçable sur les marchés du continent, avec sources, dates et statuts de validation. C'est, très concrètement, du corpus.
Encadrer	Une charte de l'intelligence artificielle en seize articles fixe nos règles internes : validation humaine obligatoire, transparence de l'assistance, traçabilité, réversibilité, non-exploitation des contenus des utilisateurs.
Enseigner	Deux des douze modules de la Méthode MAYALA® sont consacrés à l'intelligence artificielle dans la chaîne de valeur et aux droits qui s'y attachent. Un professionnel qui ne sait pas déclarer son usage de l'IA ne pourra bientôt plus défendre un dossier.
Outils	Maya OS applique ces principes dans l'environnement de travail quotidien : origine des propositions identifiable, contenus exportables, assistance désactivable.

UNE EXIGENCE QUE NOUS NOUS IMPOSONS D'ABORD

Nous demandons aux auteurs de déclarer leurs usages de l'intelligence artificielle. Il serait indéfendable de ne pas nous l'appliquer. Nos propres productions indiquent la nature de l'outil employé, les étapes concernées et la part du résultat conservée. La transparence n'a de valeur que si elle commence par soi.

« Ce qui n'est pas numérisé n'existe pas pour la machine. »

OUVERTURE

La bataille se déplace

Il y a vingt ans, la question était : qui a le droit de raconter l'Afrique ? Nous avons avancé, et cette bataille-là, nous sommes en train de la gagner — nos auteurs racontent, nos films circulent, nos regards s'imposent.

Aujourd'hui, la question est devenue : qui possède et distribue ces récits ? C'est le combat des droits, des catalogues et des circuits, et c'est celui auquel je consacre l'essentiel de mon travail.

Mais demain — et demain a déjà commencé — la question sera : avec quelles données les machines apprendront-elles à nous raconter ? Si nous laissons cette question aux autres, nous aurons gagné le droit de raconter au moment précis où ce droit cessera de suffire. Numérisons, documentons, licencions, consentons, hébergeons. Ce sont des verbes administratifs. Ils décideront pourtant de la manière dont nos petits-enfants seront représentés.

« L'avenir ne se raconte pas. Il se bâtit ensemble. »

Je ne cherche pas une place. J'installe un sujet.

TZIPPORAH MAYALA · L'ÉTOILE DE L'AFRIQUE



L'ÉTOILE DE L'AFRIQUE

L'ÉTOILE DE L'AFRIQUE

Observatoire de l'Audiovisuel Africain

Association loi 1901 · SIREN 939 011 052

52, rue du Sahel — 75012 Paris, France

contact@letoiledelafrique.fr

letoiledelafrique.fr · observatoire-africain.com

« *L'avenir ne se raconte pas. Il se bâtit ensemble.* »

AFRIQUE · EUROPE · DIASPORAS