



Treize jours pour construire une rédaction qui ne dort jamais

Comment une machine de collecte, de calcul, d'écriture, de design et de diffusion fabrique seule, chaque jour, l'information musicale de l'Afrique de l'Ouest et du monde, avec des sources gratuites, des agents qui se surveillent entre eux et zéro employé de nuit.

59 214

relevés de classements en 13 jours

23 705

titres suivis

3 205

artistes connus, dont 671 actifs

0 €

dépensés en données : uniquement des sources publiques et gratuites

Le jour où le site s'est mis à travailler sans nous

Passer de 0 à 1, ce n'est pas publier une première page. C'est le moment où un système fait quelque chose d'utile alors que personne ne le lui demande. Pour Cauri Charts, ce moment a une heure précise : 1 h du matin, heure de Lomé, quand un réveil déclenche une chaîne de vingt étapes qui collecte, calcule, écrit, illustre, programme, puis se relit elle-même.

L'Afrique de l'Ouest produit l'une des musiques les plus écoutées de la planète, mais elle n'a jamais eu de classement de référence qui lui soit propre. Les plateformes publient chacune leurs tops, dans leur coin et selon leurs règles. Les médias spécialisés existent, mais une rédaction coûte cher, et l'actualité ne s'arrête pas à 18 h.

Cauri Charts part d'une hypothèse simple : **une industrie de l'information peut se compresser en logiciel**. Un classement, c'est des relevés et une formule. Une rédaction, c'est un agenda, des sources, un style et une relecture. Un studio graphique, c'est une charte et des règles de composition. Une régie de diffusion, c'est un calendrier et des canaux. Chacun de ces métiers peut être décrit précisément, donc automatisé, à condition de ne jamais confondre automatique et approximatif.

Le premier commit date du 29 septembre 2026. Le 30 septembre, la base enregistrait ses premiers relevés. Treize jours plus tard, elle en compte 59 214, venus d'Apple Music, Shazam, Deezer et YouTube, sur 16 pays. Le site publie chaque jour un classement, des magazines, des portraits d'artistes et des posts sur Instagram, Facebook, TikTok et Telegram. Depuis le 11 octobre, il couvre aussi l'actualité musicale mondiale, deux fois par jour, avec des flashes déclenchés toutes les dix minutes si une vraie nouvelle tombe.

Cet article raconte comment cette machine est pensée : ses capteurs, sa formule, son horloge, ses agents d'intelligence artificielle, sa façon de chercher l'information, son atelier visuel et sa diffusion. Il raconte aussi ses garde-fous, parce que l'essentiel de l'ingéniosité n'est pas de produire beaucoup, mais de produire **vrai**, sans surveillance humaine permanente.

« Une machine qui publie seule n'a de valeur que si l'on peut lui faire confiance quand personne ne regarde. »

CE QUE FAIT LA MACHINE EN UNE JOURNÉE

- Relève les tops de plusieurs plateformes par pays
- Calcule quatre classements : Prime, Buzz, Fresh, Icon
- Découvre de nouveaux artistes et confirme leur pays
- Écrit Cauri Matin, Cauri Hebdo et Cauri Monde
- Cherche les records et les chiffres de streams
- Détecte les flashes et les fait confirmer
- Compose des visuels en 4:5 et 9:16
- Programme les posts dans Buffer et Telegram
- Relit tout le site et note ses propres erreurs

Cinq organes, une boucle fermée

La machine est pensée comme un organisme : des capteurs qui perçoivent, un cerveau qui calcule, un atelier qui fabrique, une voix qui diffuse et un système immunitaire qui contrôle. Chaque organe peut tomber en panne sans entraîner les autres.



Les constats du contrôle corrigent les données du lendemain : la machine apprend de ses propres erreurs

La pile technique

COUCHE	CHOIX
Application	Django (Python), gabarits Jinja
Hébergement	Vercel (fonctions + cache CDN)
Base de données	PostgreSQL serverless (Neon)
Fichiers	Vercel Blob (images des posts)
Calcul lourd	GitHub Actions (5 workflows)
Images	Pillow, rendu en 2160 × 2700 et 2160 × 3840
IA	Gemini, puis Groq, puis OpenRouter en secours
Diffusion	API Buffer, API Telegram

Trois principes de conception

1 · LE RÉEL UNIQUEMENT

Aucun chiffre n'est inventé. Une donnée manquante produit un « pas encore assez de données », jamais un remplissage. Un pays qui suit moins de 8 artistes n'a pas de classement.

2 · GRATUIT PAR CONSTRUCTION

Chaque source est publique et gratuite : tops des plateformes, API Deezer, YouTube, flux RSS, Kword. L'IA tourne sur les offres gratuites, avec un budget d'appels surveillé.

3 · TOUT EST REJOUABLE

Chaque étape est idempotente : relancer la nuit ne crée ni doublon ni second post. Une contrainte d'unicité (jour + rubrique) protège la file de publication.

Le code tient en un peu plus de 20 000 lignes de Python réparties en 29 services. Ce chiffre surprend souvent : la complexité n'est pas dans la quantité de code, mais dans la précision des règles. Chaque règle est née d'un cas réel, par exemple un homonyme sur Deezer, une pochette dont le texte se mêlait au titre, ou un post annonçant « ce soir » une émission déjà passée.

Transformer le bruit des plateformes en un seul classement

Chaque plateforme parle sa propre langue. Apple Music publie un top par pays, Shazam mesure ce que les gens cherchent, YouTube compte des vues, Deezer expose une popularité. Le premier travail de la machine est de faire parler toutes ces sources d'une seule voix.

Les importeurs relèvent les classements chaque nuit, pays par pays. Un normaliseur nettoie les titres (« feat. », « Remix », « Official Video », accents, majuscules), puis un module de rapprochement associe chaque entrée à un titre et à un artiste uniques. C'est l'étape la plus délicate : « Rema » sur une plateforme peut désigner un autre

artiste sur la suivante. La règle retenue est stricte : l'identité se confirme par plusieurs indices concordants, et le pays d'un artiste n'est retenu qu'une fois confirmé.

Autour du cœur tourne une ronde d'agents de découverte : liste de suivi, découverte d'artistes, profils (photo, nombre de fans), classification par pays, chaînes YouTube complètes, photos HD depuis Apple Music, Wikimedia Commons, YouTube et Deezer. En treize jours, la base est passée de quelques dizaines d'artistes à 3 205, dont 671 actifs.

SOURCE	RELEVÉS	CE QU'ELLE MESURE
Popularité Deezer	21 064	Popularité calculée des titres trouvés sur Deezer
Apple Music	14 800	Tops officiels par pays
Shazam	10 200	Ce que le public cherche à identifier : signal de découverte
Vues YouTube	8 179	Progression des vues des clips relevés chaque nuit
Deezer	3 800	Tops Deezer par pays
Tendances YouTube	1 080	Tendances musique par pays

La formule

Chaque apparition dans un top rapporte des points selon la position, normalisés par la taille du top : être 1^{er} d'un top 10 vaut autant qu'être 1^{er} d'un top 200.

$$\text{points} = 100 \times (\text{taille} + 1 - \text{position}) \div \text{taille} \div \text{nb_sources}(\text{pays})$$

La division par le nombre de sources du pays donne le même poids à chaque pays : un pays riche en plateformes n'écrase pas les autres.

QUATRE LECTURES DU MÊME MARCHÉ

- **Prime** : la somme des points, avec un plafond de titres par pays dans le Top régional
- **Buzz** : hors Top, les titres qui montent le plus et ceux portés par les signaux de découverte (Shazam)
- **Fresh** : les sorties de moins de 8 semaines et les artistes émergents
- **Icon** : la durée, sur toute l'histoire de la base

LES RÈGLES ANTI-TRICHE

- Un pays ne classe que ses propres artistes : un tube nigérian au Liberia ne remplit pas le Top libérien
- Sous 8 artistes suivis, pas de classement national
- Égalité départagée par le nombre de plateformes, puis par la meilleure position
- Le classement officiel hebdomadaire est un brouillon validé dans l'administration

Une horloge à trois réveils

Une machine autonome dépend de son horloge. Or aucune horloge gratuite n'est fiable à 100 %. La solution retenue : trois réveils indépendants qui se couvrent les uns les autres, et des tâches conçues pour pouvoir être relancées sans risque.



HEURE (UTC)	RÉVEIL	CE QUI SE PASSE
01:00	Vercel	Chaîne de nuit complète : relevés, calculs, articles, posts, agents de contrôle
04:23	GitHub Actions	Secours : relance la nuit uniquement si elle n'a pas tourné
05:00	Vercel	Rattrape les étapes en échec (sources indisponibles, quotas)
06:37	GitHub Actions	Second secours du rattrapage
07:00	Vercel	Cauri Monde, édition de 8 h : l'actualité musicale mondiale
19:00	Vercel	Cauri Monde, édition de 20 h : projets, records, sorties
toutes les 10 min	Réveil externe	Publie les posts arrivés à échéance, puis guette un flash
dim. 03:30	GitHub Actions	Sauvegarde hebdomadaire de la base

Pourquoi trois réveils ?

Les crons de Vercel sont précis mais limités en durée d'exécution. Ceux de GitHub Actions ont beaucoup de temps de calcul, mais partent souvent en retard. Aucun service gratuit ne propose un réveil toutes les dix minutes. La machine combine donc les trois : Vercel donne l'heure, GitHub Actions fournit la puissance, et un réveil externe, authentifié par un jeton secret, donne la fréquence.

Chaque réveil passe par un point d'entrée protégé qui ne fait que **déléguer** : il déclenche le bon workflow GitHub (édition matin, soir, flash ou file) et répond en quelques millisecondes. Le travail lourd ne consomme donc jamais le temps de calcul du site.

LA CHAÎNE DE NUIT, ÉTAPE PAR ÉTAPE

- 1 · Classements des plateformes
- 2 · Relevé des vues · 3 · Sources calculées
- 4 · Actualités
- 7 · Classement officiel (brouillon)
- 8 · Posts réseaux · 8a Cauri Matin
- 8 bis Telegram · 8 ter Cauri Hebdo
- 8 quater Portrait · 8 sexes Formats récurrents
- 10a-h · Suivi, découverte, profils, pays, Deezer, YouTube, photos HD
- 10k · Distinctions du mois · 10l Actu des artistes
- 10i · Statistiques des publications
- 10j · Agents de contrôle
- 11 · Ménage (cache, fichiers orphelins)

Chaque étape a son propre budget de temps. Une étape en échec est notée et rattrapée, sans bloquer les suivantes.

L'IA rédige, elle ne témoigne jamais

Le choix fondateur tient en une phrase : un modèle de langage n'est jamais une source. Il reformule, résume, titre et classe des faits qu'on lui fournit, mais il n'a pas le droit d'en ajouter. Toute l'architecture des agents découle de ce principe.

Ce principe vient d'un test. On a demandé à une IA, sans accès au web, les records de l'album d'un rappeur français. Elle a répondu avec assurance, avec des chiffres précis et faux. Depuis, aucun agent ne travaille « de mémoire ». Chaque agent reçoit un dossier fermé (les articles, les relevés, les chiffres calculés) et doit répondre dans un format JSON strict, champ par champ. Une réponse hors

format est rejetée ; une affirmation absente du dossier est supprimée par le relecteur.

Les agents utilisent les offres gratuites de plusieurs fournisseurs, avec une chaîne de secours : Gemini d'abord, puis Groq, puis OpenRouter. Les demandes sont groupées par lots de dix sujets pour économiser les quotas. Quand tous les quotas sont épuisés, la machine ne publie pas un texte bâclé : elle reporte l'étape au rattrapage.

RÉDACTEUR

Cauri Monde

Pour chaque sujet : rubrique, artiste, titre percutant, accroche, résumé, mot choc, question au public. Il signale aussi si le sujet est sensible ou ancien.

CHERCHEUR

Records

Interroge Google Actualités sur environ 90 artistes suivis avec des mots-clés (record, certifié, platine...). Il ne garde que les faits tirés des articles et contenant au moins un chiffre.

GUETTEUR

Flash

Toutes les 10 minutes, il cherche une nouvelle importante. Il exige au moins un grand média de référence ou trois médias concordants avant toute publication.

CALCULATEUR

Caps de streams

Sans IA : il lit les compteurs publics Kworb, calcule le prochain palier rond et annonce « bientôt » à 3 jours ou moins. Au-delà de 10 milliards, il déclenche un flash.

JOURNALISTE

Cauri Matin & Hebdo

Écrit le numéro du jour et le magazine de la semaine à partir du même classement que la page d'accueil. Un n°1 contradictoire est impossible.

BIOGRAPHE

Portraits

Assemble le parcours d'un artiste, ses meilleures positions et ses chiffres Spotify (auditeurs mensuels, pic, titres les plus écoutés).

Les cinq agents de contrôle (chaque nuit, après la publication)

AGENT	QUESTION POSÉE	ACTION
Éditions	Cauri Matin et Cauri Hebdo donnent-ils le même n°1 que le classement ?	Corrige
Catalogue	Les artistes du Top 50 ont-ils un pays confirmé, une photo, une chaîne YouTube ?	Correction ciblée
Publications	Les posts programmés ne montrent-ils que des artistes confirmés, avec des images en ligne ?	Bloque ou alerte
Relecteur IA	La légende affirme-t-elle quelque chose qui n'est pas dans ses faits ?	Alerte
Pages	Les pages répondent-elles, sans « None » ni image cassée ?	Alerte

Les agents ne se font pas confiance entre eux. Le rédacteur écrit, le relecteur doute, le guetteur exige des preuves : c'est une rédaction en miniature, avec ses contre-pouvoirs.

Lire 9 rédactions à la fois, n'en garder que le vrai

Les fans voulaient l'actualité mondiale. Le défi était de la couvrir sans rédaction, sans inventer, et sans changer d'identité : Cauri Charts reste « la référence de la musique ouest-africaine », qui regarde aussi le monde.

1 • Collecter

La machine lit les flux RSS de neuf médias musicaux et culturels de référence (Billboard, Rolling Stone, Variety, Pitchfork, NME, Consequence, HipHopDX, Music Business Worldwide, RFI). Elle ajoute une veille Google Actualités ciblée sur une liste de stars mondiales et de géants africains.

2 • Regrouper

Une même nouvelle apparaît dans plusieurs médias sous des titres différents. Les articles sont regroupés par noms propres et mots significatifs communs. Un groupe = un sujet. Le nombre de médias dans un groupe devient un indice de fiabilité et d'importance.

3 • Classer

Les sujets sont classés par audience des artistes concernés (nombre de fans Deezer) et par couverture médiatique. Règle éditoriale : **une info par artiste et par édition**, et des visages connus en priorité.

4 • Filtrer

Une nouvelle de plus de 30 jours est écartée. Une version non officielle d'un clip est refusée. Un sujet que l'IA marque comme « sensible » ou « ancien » sort de l'édition. Chaque sujet publié est inscrit dans une mémoire pour ne jamais être publié deux fois.



Une couverture Cauri Monde générée automatiquement : vraie photo de l'artiste, titre géant, zone de texte floutée pour rester lisible.

UNE INFO = UN POST = UN VISAGE

Chaque édition publie au plus 4 sujets, chacun dans son propre post, espacés d'au moins 45 minutes. La première image doit dire exactement de quoi on parle : si le nom de l'artiste n'est pas dans le titre, sa photo n'est pas utilisée.

LE FLASH : DU SIGNAL À LA PUBLICATION

T + 0	Le réveil externe déclenche le guetteur. La file publie d'abord les posts arrivés à échéance.
T + 1 min	Le guetteur relit les flux, regroupe et cherche un sujet majeur qui n'a pas encore été vu.
T + 2 min	Confirmation : au moins un grand média de référence ou trois sources. Sinon, rien ne part.
T + 4 min	Rédaction, couverture « choc » en 4:5 et 9:16, archivage sur le site, envoi vers Buffer et Telegram.

Un directeur artistique écrit en Python

Un post doit arrêter le pouce en moins d'une seconde. Pour cela, l'atelier compose chaque image à partir de règles qu'un graphiste appliquerait : hiérarchie, contraste, respiration, et surtout ne jamais cacher un visage.

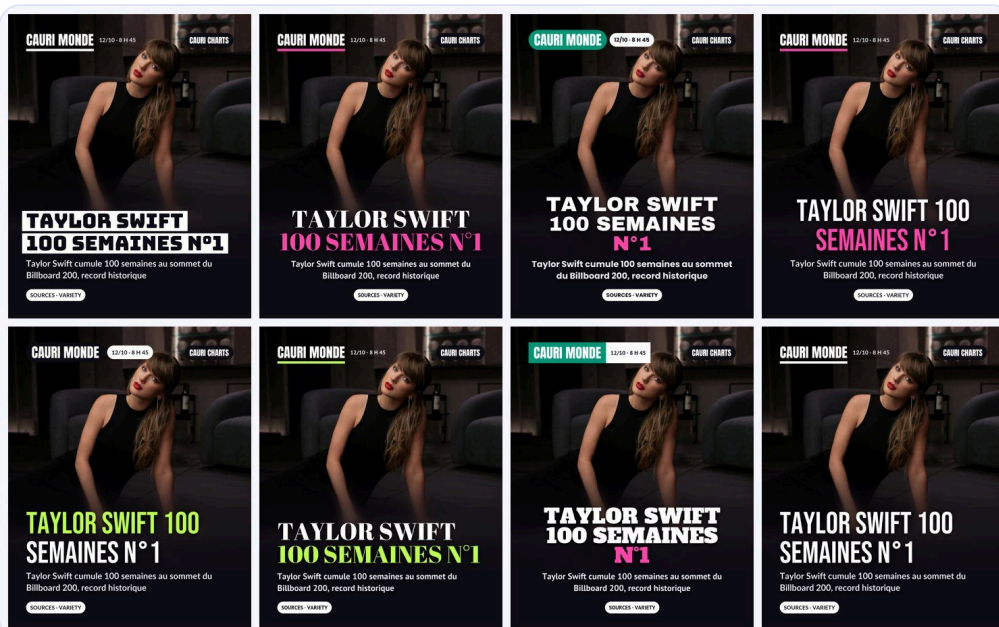


Planche de styles : le même moteur, des combinaisons différentes. Chaque post tire son style d'une « graine » (son nom unique), et toutes les images d'un carrousel partagent ce style.

Le moteur de styles

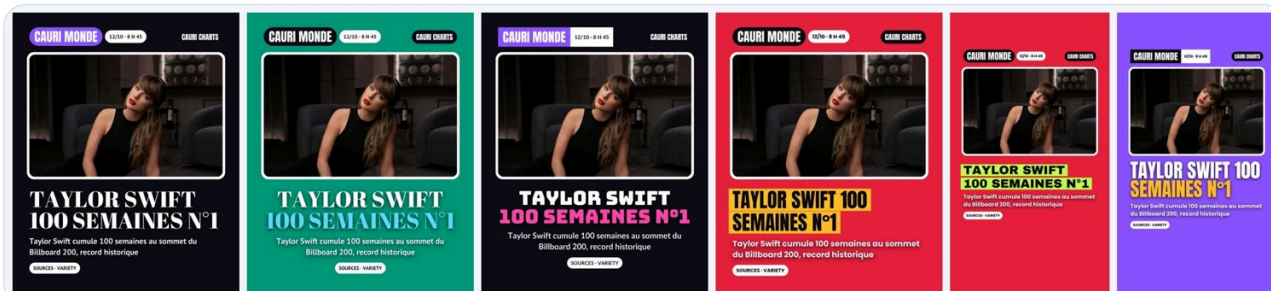
Pour chaque post, le moteur choisit une police parmi six (Anton, Bebas Neue, Archivo Black, Alfa Slab One, Bungee, Abril Fatface), une couleur d'accent parmi six, un traitement du titre (dernière ligne colorée, première ligne, surligné, alterné), une mise en page (texte en bas, en haut, ou affiche avec photo encadrée), un alignement et un style de bandeau.

$$6 \times 6 \times 4 \times 3 \times 2 \times 3 = 2\,592$$

combinaisons possibles, tirées de façon déterministe : un post régénéré garde le même style.

Les règles de composition

- **Budget de texte** : le titre reste dans le tiers bas (ou haut) de l'image, pour ne jamais couvrir un visage.
- **Zone floutée et assombrie** sous le texte, plus une ombre portée : lisible sur une pochette blanche comme sur une photo de nuit.
- **Coupe propre** des titres trop longs, sans couper un mot ni un caractère spécial.
- **Jamais de fond uni** : photo de l'artiste (notre base d'abord, puis Deezer seulement si le nom correspond exactement), objet du sujet (trophée, vinyle, micro, domaine public) ou scène de concert.
- **Rendu en double définition** (2160 px de large) pour les formats feed 4:5 et story 9:16.



Mise en page « affiche » : fond de couleur, photo encadrée, texte sous le cadre.

Une grille de programmes **comme une chaîne de télé**

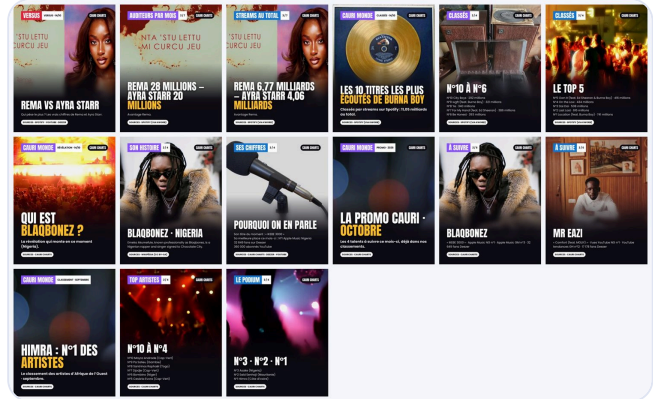
Publier n'est pas « poster quand c'est prêt ». La machine tient une grille : des rendez-vous fixes que le public apprend à attendre, et des posts d'actualité qui s'y glissent.

La grille hebdomadaire

QUAND	FORMAT
Chaque matin	Cauri Matin : le n°1 du jour et l'annonce de la soirée
8 h · 20 h	Cauri Monde : jusqu'à 4 posts par édition
Lundi 21 h	10 titres classés
Mercredi 21 h	Versus : deux artistes face à face, le public vote
Vendredi 21 h	Qui est X ? Portrait d'un artiste
1 ^{er} du mois	Classement mensuel des artistes
1 ^{er} lundi	Promo Cauri
Semaine	Cauri Hebdo : magazine et carrousel
À tout moment	Flash, si une vraie nouvelle tombe

Une question dans chaque post

Chaque post Cauri Monde se termine par une question au public, écrite par l'agent rédacteur à partir du sujet. Les commentaires deviennent un signal : ils alimentent les votes de l'émission mensuelle.



Les formats récurrents : versus, 10 titres classés, qui est X, classement mensuel, promo.

Le circuit d'un post

- Les images sont rendues, puis envoyées sur le stockage Blob sous un **nom unique horodaté**, pour qu'aucun cache ne serve une ancienne version.
- Le post est enregistré dans la base (une seule ligne par jour et par rubrique) avec son heure de publication.
- À l'heure dite, la file l'envoie à Buffer, qui publie sur Instagram et Facebook.
- TikTok en mode notification** : le fondateur reçoit le post sur son téléphone et ajoute un son tendance avant de publier. C'est le seul geste humain, et il est voulu : le son tendance fait la portée sur TikTok.
- En parallèle, la chaîne Telegram reçoit le classement et les éditions.
- Chaque article est aussi archivé sur le site, dans sa série (Cauri Matin, Cauri Hebdo, Cauri Monde), avec la couverture du post comme image.

LE SITE COMME VITRINE PERMANENTE

Pages des classements (Prime, Buzz, Fresh, Icon, par pays et par plateforme), pages des séries, et la rubrique « Les plus écoutés dans le monde » alimentée par les compteurs publics de Spotify. Toutes ces pages sont servies depuis le cache CDN : un visiteur de plus ne coûte presque rien.

Quand la machine passe à l'antenne

Le même cerveau qui produit les posts alimente The Monthly Show, la remise des prix mensuelle diffusée en direct comme une émission de télévision. Une antenne graphique (Remotion dans le navigateur, capturée par OBS) affiche les classements, les nominés et les clips officiels. Une régie pilote le conducteur. Un collecteur lit en temps réel les commentaires du direct TikTok pour compter les votes du public.

Le son suit les mêmes règles que les images : des ambiances tirées de banques sonores réelles, des musiques du domaine public, et aucun effet bon marché.

CE QUE LES PANNES NOUS ONT APPRIS

- **Le cache est un organe vital.** Un simple en-tête Vary: Cookie empêchait le CDN de servir les pages : chaque visite relançait Django et le quota de calcul a été épuisé. Corrigé par un middleware qui purge cet en-tête sur les pages publiques. Le cache répond désormais.
- **Un nom de fichier est une promesse.** Réutiliser un nom d'image faisait publier l'ancienne version, servie par les caches. Tous les fichiers sont désormais horodatés.
- **Une seule source de vérité.** Cauri Matin et la page d'accueil calculaient deux n°1 différents. Ils lisent maintenant le même classement, et un agent le vérifie chaque nuit.
- **L'IA sans dossier invente.** D'où la règle : elle résume, elle ne raconte pas.

Ce que « de 0 à 1 » veut vraiment dire ici

L'innovation de Cauri Charts n'est pas une technologie isolée. Ni Django, ni Pillow, ni les modèles de langage ne sont nouveaux. Ce qui est nouveau, c'est l'**assemblage** : une agence de données, une rédaction, un studio graphique, une régie de diffusion et un service qualité, compressés en un seul système qui se réveille seul, travaille seul, se relit seul et ne coûte presque rien en données. C'est cela, passer de 0 à 1 : créer une catégorie qui n'existait pas, une **industrie de l'information musicale autonome** pour une région que les grands classements ignoraient.

16

pays dans la base

~20

étapes dans la chaîne de nuit

144

réveils flash par jour

1

geste humain : le son TikTok

La suite

Votes Instagram en direct, antenne verticale pour le direct mobile, émission complète de près de quatre heures, moteur de styles étendu à toutes les rubriques. Objectif : 10 000 abonnés d'ici fin décembre 2026. La machine, elle, a déjà commencé : pendant que vous lisiez cet article, elle a probablement vérifié deux fois s'il y avait un flash.

La musique ouest-africaine avait ses stars. Elle a maintenant sa machine.