



FICHE MÉMO IMPRIMABLE

Décrypter la consommation électrique d'une IA

Une fiche pratique pour comprendre ce qui consomme réellement de l'électricité lorsqu'un modèle d'intelligence artificielle est entraîné ou utilisé. À garder sous la main pour éviter les comparaisons trompeuses.

Au sommaire

- 1 Les 5 repères essentiels
- 2 Ne pas confondre les notions
- 3 Avant de croire une estimation
- 4 Suivre le parcours de l'électricité
- 5 La règle de comparaison

Les 5 repères essentiels

Il n'existe pas de consommation électrique universelle pour une requête d'IA : le modèle, la tâche, la longueur du traitement et le matériel font varier le résultat.

La consommation ne se limite pas aux puces de calcul : mémoire, stockage, réseau, alimentation et refroidissement participent aussi au bilan.

Une baisse de l'énergie consommée par tâche ne garantit pas une baisse totale si le nombre de demandes augmente plus vite que les gains d'efficacité.

L'entraînement mobilise des calculs nombreux et prolongés ; l'inférence correspond à l'utilisation répétée d'un modèle déjà entraîné.

La puissance se mesure comme un débit instantané en watts ; l'énergie correspond à ce débit sur une durée, en wattheures ou kilowattheures.

Ne pas confondre les notions

Notion	Ce qu'elle indique	Question à poser
Puissance	Le débit instantané d'électricité appelé par un équipement ou une installation.	À quel moment et avec quelle intensité les équipements fonctionnent-ils ?
Énergie	La quantité d'électricité consommée pendant une durée.	Combien de temps le traitement a-t-il mobilisé les équipements ?
Périmètre de mesure	Les éléments inclus dans l'estimation : accélérateur seul, serveur ou centre de données.	Le calcul, le refroidissement et les équipements auxiliaires sont-ils comptés ?
Impact environnemental	Une dimension particulière parmi l'électricité, les émissions, l'eau ou les matériaux.	Quel indicateur est réellement mesuré ?

Avant de croire une estimation

Identifier le modèle utilisé

Un modèle à l'autre, les besoins de calcul peuvent différer.

Préciser le type de tâche

Texte court, analyse de documents, raisonnement en plusieurs étapes ou génération d'image ne mobilisent pas les mêmes ressources.

Vérifier la longueur du traitement

La longueur de la demande et celle de la réponse influencent la consommation.

Distinguer entraînement et inférence

Il s'agit de phases différentes, avec des usages et des durées de calcul différents.

Regarder le matériel mobilisé

Accélérateurs, processeurs, mémoire, stockage et réseau peuvent tous participer au fonctionnement.

Contrôler le périmètre mesuré

Une mesure limitée à l'accélérateur ne couvre pas nécessairement l'ensemble du serveur ou du site.

Vérifier l'indicateur annoncé

Électricité, émissions de gaz à effet de serre, eau et impacts matériels ne décrivent pas la même réalité.

CHAÎNE ÉLECTRIQUE D'UN SYSTÈME D'INTELLIGENCE ARTIFICIELLE

Du traitement des données au résultat : le parcours de l'énergie et de l'information dans une infrastructure IA

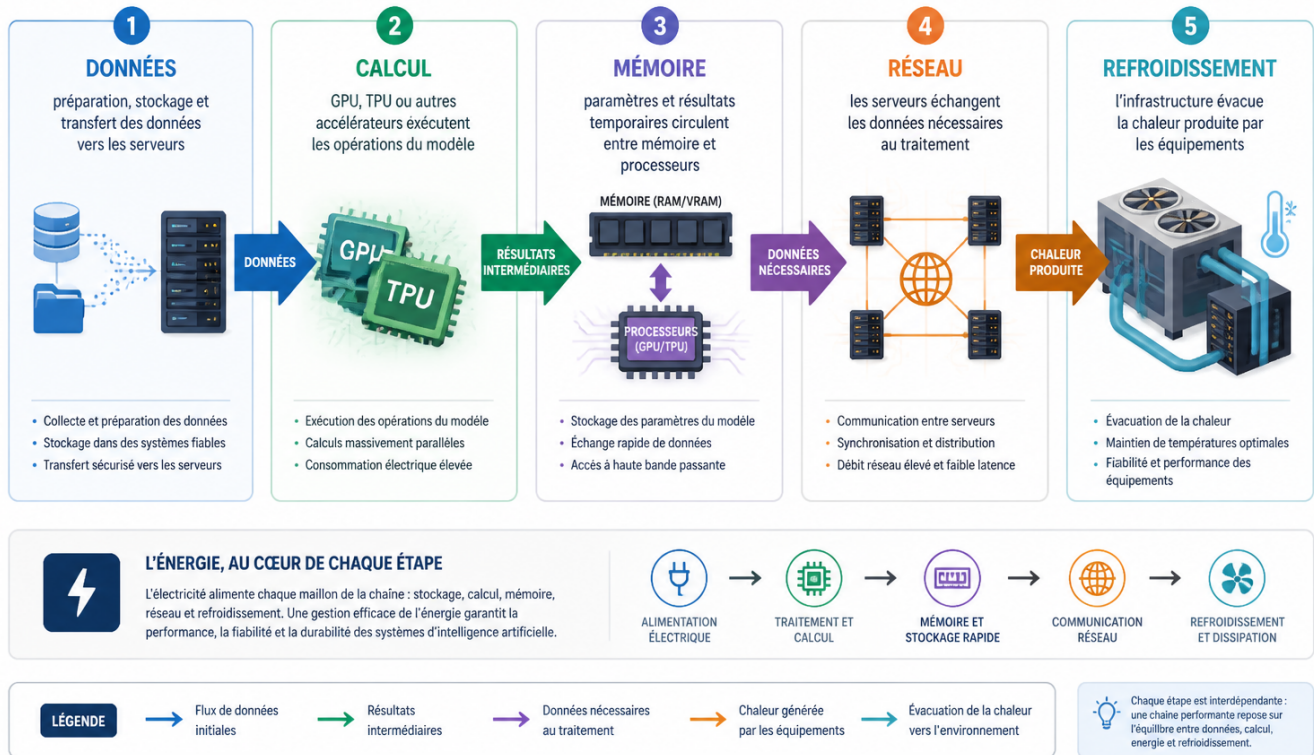
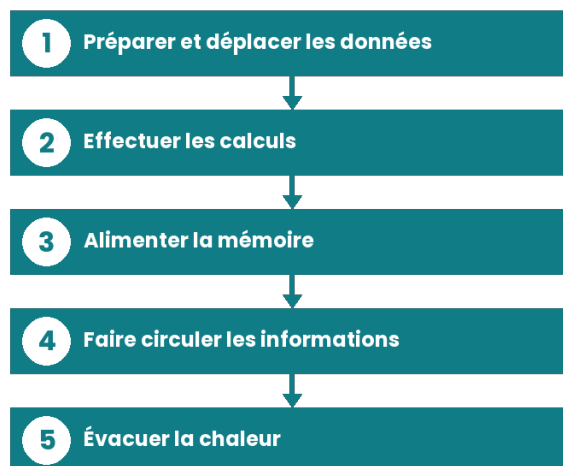


SCHÉMA Le parcours de la consommation électrique d'une IA : des données au calcul, puis à la mémoire, au réseau et au refroidissement.

Suivre le parcours de l'électricité



1 **Préparer et déplacer les données**

Les données sont préparées, stockées et transférées avant ou pendant le traitement.

2 **Effectuer les calculs**

Des GPU, TPU ou autres accélérateurs réalisent de nombreuses opérations en parallèle, avec l'appui de processeurs classiques.

3 **Alimenter la mémoire**

Les paramètres du modèle et les résultats temporaires doivent être conservés et déplacés en mémoire.

4 **Faire circuler les informations**

Les serveurs échangent les données grâce aux réseaux internes et aux autres équipements de communication.

5 **Évacuer la chaleur**

Le refroidissement maintient l'infrastructure en fonctionnement et ajoute une consommation au calcul informatique.

! Un chiffre à remettre dans son contexte

L'Agence internationale de l'énergie indique que les centres de données représentaient environ 1 % de la demande mondiale d'électricité en 2022. Cette valeur concerne l'ensemble des centres de données, et non l'intelligence artificielle seule : elle ne permet donc pas d'isoler directement la part des modèles d'IA.

La règle de comparaison

Pour comparer deux usages de l'intelligence artificielle, alignez au minimum le modèle, la tâche, la longueur du traitement, le matériel et le périmètre de mesure. Sans ces précisions, deux chiffres apparemment comparables peuvent en réalité mesurer des réalités différentes.

Les estimations de consommation doivent être interprétées selon leur périmètre et leurs hypothèses ; l'électricité, les émissions, l'eau et les impacts matériels sont des indicateurs distincts.