

Exploration de 3 outils de la plateforme IA de Vitasciences !

La [plateforme IA de vitasciences](#) est recommandée par l'inspection générale et a pour avantage de ne pas nécessiter d'entrer des données personnelles. Dans cet article vous sont présentés trois outils de cette plateforme qui permettent de mieux comprendre le fonctionnement des modèles d'IA. Bonne lecture !

Audrey Boisson – IAN SVT Ac-Guyane

● OUTILS 1 : « Génération Texte »

➤ Description technique de l'utilité de l'outil :



L'outil « [Génération Texte](#) » permet d'interagir avec différents modèles de langage comme Mixtral, GPT-4o-mini, Llama-3, ... et de découvrir leur fonctionnement. Dans l'interface utilisateur il suffit d'entrer un prompt (consigne textuelle), de choisir un modèle de langage ainsi qu'un niveau de créativité pour générer en temps réel une réponse textuelle. Ce niveau de créativité dans les réponses ajuste la « température » du modèle de langage, c'est-à-dire le degré d'aléatoire dans la génération du texte. L'interface propose également un jeu de « pré-prompts » qui permettent de conditionner le modèle pour un usage particulier (ex : création de quiz, générateur de poème, ...)

➤ Prise en main de l'outil :

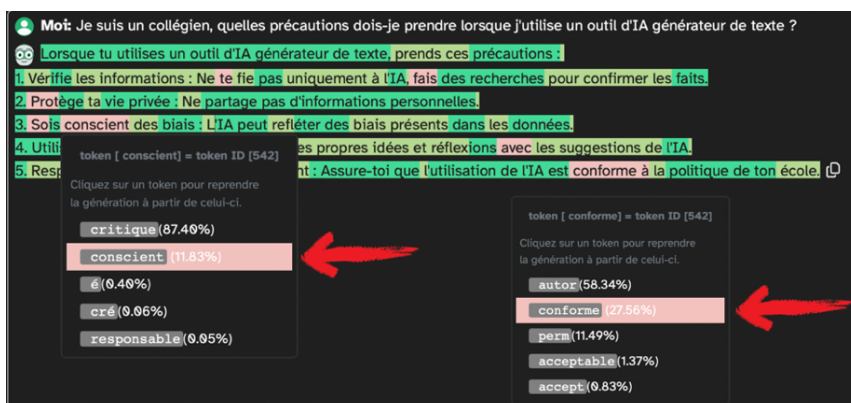


Étape 1 :

Commencez par **choisir un modèle de langage** (ex : GPT-4o-mini).

Pour se familiariser avec l'outil commencez par **saisir un prompt** simple (ex : « Je suis un collégien, quelles précautions dois-je prendre lorsque j'utilise un outil d'IA générateur de texte »).

Affinez progressivement ce prompt afin de tester la pertinence et la diversité des réponses générées.



Étape 2 :

Expérimentez en modifiant les ***tokens** :

** Les tokens sont les unités de texte traitées par le modèle de langage, le modèle va alors prédire chaque mot en fonction du contexte précédent.*

En modifiant un token la réponse est entièrement reprise est modifiée en fonction de ce nouveau token.

Étape 3

Expérimentez en utilisant la fonction de **réglage du niveau de créativité** (caractère aléatoire). Nous pouvons alors observer comment la réponse générée varie en fonction de ce paramètre.

Plus le niveau de créativité se rapproche de 0% plus les réponses sont factuelles et plus le modèle devient déterministe et répétitif. A l'extrême inverse lorsqu'il tend à se rapprocher de 100% les réponses sont de plus en plus imaginatives, parfois inattendues, et nous pouvons observer les premières « hallucinations » de l'IA.



➤ Défis rencontrés lors de l'utilisation de l'outil :

- Un premier défi est d'obtenir des réponses précises à des questions complexes, car elles sont souvent trop générales et peu adaptées au niveau de langage de l'utilisateur.
 - ⇒ Stratégies de résolution : reformulez et affinez le prompt en précisant le contexte, la tonalité et le niveau de développement souhaité dans la réponse.
- Certaines réponses à des questions plus scientifiques peuvent paraître incohérentes voire contradictoires (ex : données relatives au niveau de précipitation en Guyane Française).
 - ⇒ Stratégies de résolution : formulez des consignes plus courtes et ciblées, tout en croisant les réponses générées avec d'autres sources afin d'en vérifier la validité.

➤ Apports et limites de l'outils :

- L'interface utilisateur est assez intuitive, l'outil est accessible tant pour des élèves que pour des adultes et permet d'expérimenter avec des modèles de langage avancés.
- Les possibilités de régler le niveau de créativité et de modifier spécifiquement certains tokens sont intéressants dans la compréhension du fonctionnement général du modèle.
- Les modèles de langage disponibles ont cependant une capacité limitée pour traiter des prompts longs. L'IA peut également générer des informations erronées avec l'utilisation notamment du niveau de créativité du modèle, ce qui nécessite un esprit critique de l'utilisateur.

Remarque : Il est important de chercher à **exercer un regard critique sur ce que produit le modèle de langage**, en essayant de soulever les biais et les erreurs potentielles que l'outil peut procurer.

● OUTILS 2 : « Entraînement Images »

➤ Description technique de l'utilité de l'outil :



Comme tous les outils proposés par vittasciences, l'interface « [Entraînement Image](#) » fonctionne sans installation et sur tout support. Elle permet d'entraîner un modèle d'IA à reconnaître différentes catégories d'images, en important ses propres données (importées depuis l'ordinateur ou avec la webcam) ou en utilisant un jeu de données libres de droits.

L'interface respecte les RGPD, les droits d'auteur et les droits à l'image. Elle possède 4 parties :

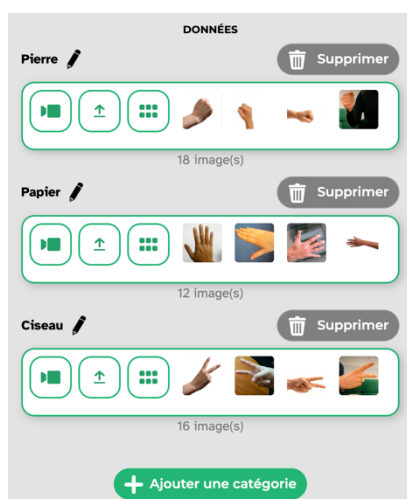
- Données : importation de données d'entraînement.
- Entraînement du modèle.
- Aperçu : elle permet de tester le modèle et de comprendre son fonctionnement.
- Utilisation : elle permet d'utiliser le modèle en dehors de l'interface (au choix : adacraft pour interagir avec des microcontrôleurs ; python ; téléchargement du modèle entraîné sur le PC).

➤ Prise en main de l'outil :

Étape 1 :

Essayez par exemple de différencier une main qui joue à pierre-feuille-ciseaux (jeu du chifoumi). Pour cela commencez par créer 3 catégories (C1 : pierre ; C2 : feuille ; C3 : ciseaux) dans la partie « données » de l'interface.

Utilisez la banque d'images libres de droit de l'interface pour importer des données d'entraînement dans chacune des catégories créées.



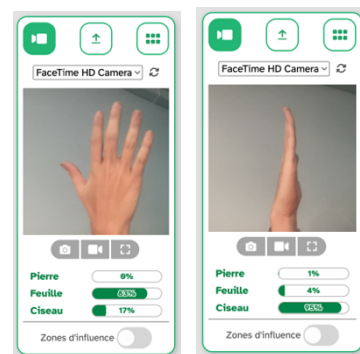
Étape 2 :

Une fois les données importées, lancez l'entraînement du modèle.

Étape 3 :

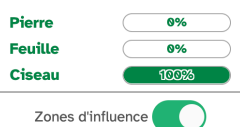
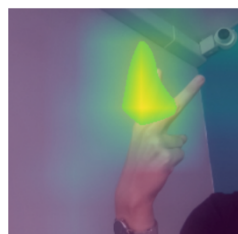
Testez ensuite le modèle dans la partie aperçu avec votre webcam, en lui présentant votre main (pierre, feuille ou ciseau). La prédiction générée par l'IA est actualisée plusieurs fois par secondes !

Observations : quand vous positionnez votre main de profil devant la caméra pour représenter une feuille, le modèle prédit à environ 90-95% que cette dernière appartient à la catégorie « ciseaux ».

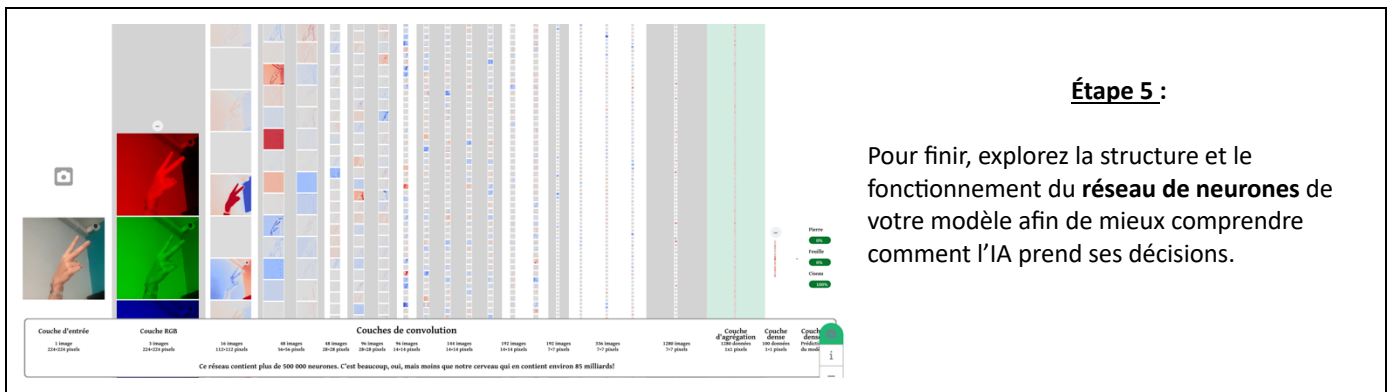


Afin d'améliorer votre modèle, vous pouvez enrichir votre base de données d'entraînement en capturant des images de vos mains sous plusieurs angles. Vous pourrez alors observer que les erreurs dans les prédictions sont moins importantes.

Étape 4 :



Activez la **visualisation des données d'interaction** qui permettent au modèle de trouver un résultat. Cette fonctionnalité permet de mettre en évidence les zones de l'image qui sont les plus déterminantes dans les prédictions de l'IA (zones d'influence).



➤ Défis rencontrés lors de l'utilisation de l'outil :

- Un premier défi est de rendre le modèle le plus fiable possible avec relativement peu d'images.
 - ⇒ Stratégies de résolution : afin d'améliorer la robustesse de votre modèle, vous pouvez d'abord augmenter le nombre d'images par catégorie avec la banque d'images libres de droits en ligne. Ensuite, vous pouvez également effectuer des photographies de vos mains sous des angles qui n'étaient pas présents dans les clichés de la plateforme. Après un nouvel entraînement du modèle, vous pourrez observer que qu'il fait beaucoup moins d'erreurs !
- Un autre défi est de comprendre certaines erreurs de classification qui sont encore présentes malgré l'enrichissement de la base de données d'images.
 - ⇒ Stratégies de résolution : la fonctionnalité « zone d'influence » de la plateforme permet de comprendre les défaillances du modèle que vous aurez pu constater. Les zones de l'image les plus déterminantes dans les prédictions de l'IA paraissent être le haut de la main/les doigts. La prédiction « ciseaux » se rapproche de 100% quand les deux doigts sont bien dissociés et repérables facilement. Pour les deux autres catégories le nombre d'erreurs est plus important car les zones de la main sont plus difficiles à distinguer.

➤ Apports et limites des outils :

- Cet outil est assez facile à prendre en main pour créer et tester des modèles de reconnaissance d'images.
- Il permet d'importer ses propres données ou bien d'utiliser une banque d'images libres de droits, ce qui offre plusieurs possibilités intéressantes à l'utilisateur, tout en respectant les normes RGPD et les droits à l'image.
- La visualisation des zones d'influence et l'exploration de la structure et du fonctionnement du réseau de neurones du modèle favorise la compréhension du processus général de reconnaissance d'images.
- Cependant les performances du modèle dépendent beaucoup de la quantité et de la diversité des images fournies, ce qui peut nécessiter un apport important d'images de la part de l'utilisateur.
- L'outil propose également assez peu d'options avancées pour affiner le modèle, ce qui limite son usage à un cadre davantage éducatif et expérimental.

• OUTILS 3 : « Génération Image »

➤ Description technique de l'utilité de l'outil :



L'interface de [Génération d'images](#) de Vittasciences permet d'interagir avec des modèles IA images en utilisant des instructions et en modifiant les paramètres.

L'objectif de cette IA est ici de créer des visuels à partir de description textuelle en utilisant des modèles d'apprentissage. L'interface fonctionne sans installation et sur tout support.

L'interface de génération d'images de Vittasciences est composée de **4 parties** :

1. Une **zone de texte** qui permet d'interagir avec l'IA en décrivant ce que nous souhaitons obtenir sur l'image, avec possibilité d'ajouter un prompt négatif pour exclure certains éléments indésirables.
2. Une **zone de paramétrage** qui permet d'ajuster l'IA selon nos besoins (rapport d'aspect, style, ...).
3. Une **zone de visualisation des étapes d'assemblage des pixels** par l'IA qui permet d'obtenir l'image.
4. Une **zone de regroupement des images** générées par l'utilisateur ou par la communauté.

➤ Prise en main de l'outil :

Étape 1 :

Pour un premier test avec l'interface vous pouvez utiliser (par exemple) le modèle gratuit Stable-Diffusion-XL.

Demandez à l'IA de créer un visuel à partir d'un premier **prompt simple** : « *Des enfants qui font une expérience dans un cours de sciences avec un microscope.* »

→ L'image générée est floue par endroits, certaines proportions sont étranges (visages des enfants, mains, ...) et des éléments inattendus sont présents comme de nombreux objets sur la table qui ne sont pas tous identifiables, d'autres enfants à l'arrière-plan qui semblent se fondre avec le mur situé au fond de la salle, ...

Vous pouvez observer le **processus d'assemblage des pixels de manière aléatoire par l'IA**, permettant de générer une image cohérente en passant par plusieurs étapes (20 étapes sont visibles sur la plateforme).

Étape 2 :

Afin d'améliorer la qualité de l'image générée, vous pouvez enrichir votre prompt en précisant le **style** que vous souhaitez (ex : photographique). Vous pouvez également utiliser la fonction **prompt négatif** pour exclure certains éléments gênants (ex : « pas d'autres éléments sur la table » ; « pas de détails flous »).

En conservant le même prompt initial, le même prompt négatif ainsi que le même style et rapport d'aspect, vous pouvez tester de générer plusieurs fois l'image : les résultats correspondent à la description, cependant la position du microscope et des enfants et objets sur l'image varie selon les essais.



Étape 3

Conservez les paramètres initiaux et testez cette fois-ci de modifier le modèle de génération d'image utilisé (ex : SSD-AB). Ce nouveau modèle est une version réduite du modèle de diffusion stable XL (SDXL), qui offre une accélération de 60% tout en conservant des capacités de génération de texte à image de haute qualité.

→ Le contenu généré semble se rapprocher beaucoup du modèle précédemment utilisé, les étapes observables permettant d'arriver à l'image sont cependant différentes.

Une fois l'image souhaitée générée et correspondant à vos attentes, vous pouvez si vous le souhaitez l'exporter. N'oubliez pas de notifier « image générée par l'IA » en-dessous de celle-ci si vous souhaitez intégrer cette image dans un support ou l'a partager.

➤ Défis rencontrés lors de l'utilisation de l'outil :

- Un premier défi est d'obtenir une image correspondant parfaitement à votre vision.
 - ⇒ Stratégies de résolution : pour vous en rapprocher, vous pouvez affiner votre prompt en insistant sur les détails et en ajoutant certains mots clés comme le style souhaité.
Des éléments inattendus et non désirés peuvent parfois apparaître : utilisez le prompt négatif afin d'exclure spécifiquement ces éléments.
- Un autre défi est de bien cerner toutes les limites du modèle. Ce dernier a par exemple du mal à générer des mains humaines réalistes, ou encore des textes lisibles sur les images.
 - ⇒ Stratégies de résolution : tenir compte de ces contraintes pour adapter et orienter vos demandes.

➤ Apports et limites de l'outils :

- L'outil est ludique, simple à utiliser et ne nécessite pas la création de compte utilisateur ce qui est très intéressant pour l'utiliser avec des élèves ou en formation d'enseignants.
- La visualisation des différentes étapes de génération de l'image avec l'assemblage des pixels de manière aléatoire permet de mieux comprendre le processus de génération qui n'est pas un processus déterministe, un même prompt pouvant générer des résultats très différents.
- La possibilité d'ajouter un prompt négatif est réel un atout pour obtenir des images plus précises et correspondant mieux aux attentes de l'utilisateur.
- La question de la provenance des données d'entraînement n'est cependant pas abordée, ce qui pose la question du droit d'auteur !
- Les images générées peuvent également comporter des incohérences ou des artefacts malgré le prompt négatif, notamment sur des détails complexes (ex : main humaine).
- Enfin, le nombre de caractères est assez limité, notamment pour le prompt négatif (100 caractères maximum), ce qui nécessite une certaine capacité de synthèse de la part de l'utilisateur dans la rédaction du prompt.