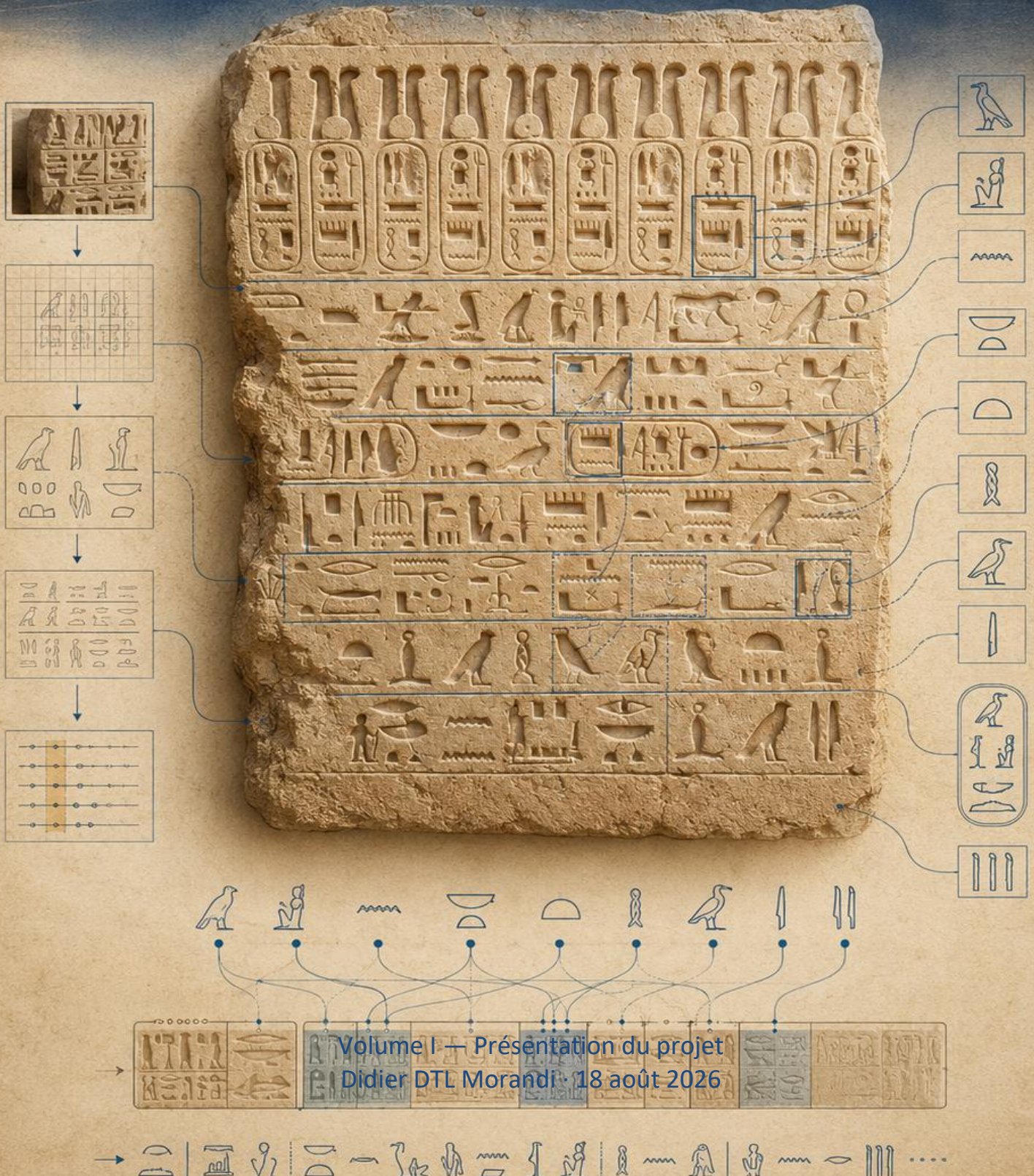


PROJET LORACRAFT

De la photographie d'une inscription à une traduction argumentée



À Bernard Mathieu,
*égyptologue de renommée mondiale et
enseignant hors pair, qui m'a transmis
bien davantage que le sens des signes,
et m'a si bien guidé lorsque le chemin
devenait incertain. Merci.*



Montpellier, le 18 août 2026.

01 Un livre pour comprendre le projet

Le premier document¹ publié sur le projet **Loracraft** est une architecture technique de référence. Il décrit avec précision les objets de données, les interfaces, les formats et les règles qui permettent à plusieurs logiciels de travailler ensemble. Cette précision est utile aux développeurs, mais elle répond mal à une question plus simple : qu'essayons-nous réellement de faire avec une photographie couverte de hiéroglyphes ?



¹ Architecture scientifique de **Loracraft**, Didier DTL Morandi, 17 août 2026.

Le bas-relief reproduit à la page précédente appartient à la Chapelle blanche du pharaon Sésostri I^{er} (XII^e dynastie), aujourd'hui reconstruite dans le musée en plein air de Karnak, en Égypte. À gauche, Sésostri I^{er}, coiffé de la couronne rouge de Basse-Égypte, fait face au dieu Amon², principale divinité du sanctuaire de Karnak. Amon présente au nez du roi un groupe composé des signes ânkh et djed : S34-R11, « vie et stabilité ». Dans leur main abaissée, le dieu et le pharaon tiennent chacun un ânkh simple. La scène représente un don divin : Amon accorde ou renouvelle au roi la vie et la stabilité. Les figures, les cartouches royaux et les inscriptions sont minutieusement sculptés en relief levé, selon la technique particulièrement raffinée qui caractérise la Chapelle blanche. © Google AI.

« Qu'essayons-nous réellement de faire avec une photographie couverte de hiéroglyphes ? ». Ce document répond à la question. Il suit une inscription depuis son entrée dans le logiciel jusqu'à sa traduction. Il explique les difficultés au moment où elles apparaissent. Lorsqu'un terme spécialisé est indispensable, il est défini dans la phrase. Les détails de code restent dans le volume technique.

EN CLAIR Loracraft n'est pas un logiciel qui regarde une stèle et annonce aussitôt ce qu'elle dit. C'est un atelier numérique : il prépare l'image, découpe la composition, propose des identifications, conserve les désaccords et aide le chercheur à construire une lecture vérifiable.

Le fil conducteur est fourni par cinq images de test (en annexe). Certaines sont faciles, d'autres dégradées, verticales ou photographiées de travers. Elles ne forment pas encore un corpus scientifique représentatif ; elles servent à découvrir les erreurs du prototype et à empêcher qu'une amélioration en casse une autre.

L'image de couverture et le texte de ce document sont une production © OpenAI ChatGPT 5.6 revue et corrigée par l'auteur (notamment toutes les notes de bas de pages).

² Amon est le dieu tutélaire de Thèbes et la principale divinité du temple de Karnak. Il ne s'agit pas du dieu Min, contrairement à l'identification fournie par Google IA.

02 Deux volumes, deux lecteurs

Le présent volume

Il est destiné au lecteur cultivé qui ne passe pas son temps à programmer en Python et qui ne pratique pas tous les jours l'épigraphie égyptienne. On y parle ici de photographie, de signes, de mots, de preuves et de décisions. Le but est de comprendre le parcours d'une inscription et l'état réel du prototype.

Le volume technique

En revanche, le document ***Architecture scientifique de Loracraft*** reste la référence des développeurs. On y conserve les schémas de données, les identifiants uniques, les interfaces de programmation, les formats JSON, les règles de versionnement et les contrats précis entre modules. Ce volume n'est pas supprimé : il devient le mode d'emploi des machines et des contributeurs techniques.

POURQUOI EST-CE IMPORTANT ? Mélanger les deux niveaux produit soit un texte vague qui ne permet pas de construire le logiciel, soit un texte exact mais illisible pour son commanditaire. Les séparer permet de garder la rigueur sans condamner le lecteur à traduire chaque paragraphe ou à lire trois cents notes de bas de page.

Les deux livres se répondent. Lorsqu'une décision de conception doit être mise en œuvre, on consulte l'architecture technique. Lorsqu'il faut expliquer le projet, décider d'une priorité ou présenter un résultat, on revient ici.

03 La chaîne complète en dix gestes

1. Recevoir la photographie et conserver l'original intact.
2. Créer une copie de travail, puis corriger éventuellement son orientation et son contraste.
3. Repérer la surface inscrite, les registres, les lignes ou colonnes, puis les cadrats.
4. Déterminer dans quel sens les signes doivent être parcourus, horizontalement ou verticalement.
5. Isoler les objets visuels et proposer, lorsque c'est possible, des codes de la liste de Gardiner.
6. Soumettre les propositions fragiles à une revue humaine.
7. Représenter les signes et leurs groupements avec le Manuel de Codage.
8. Translittérer, découper les mots et proposer des analyses grammaticales.
9. Chercher des attestations et des passages parallèles dans des corpus, lorsque les accès le permettent.
10. Construire puis publier une traduction reliée à ses preuves et à ses variantes.

RÈGLE DU JEU Une étape peut répondre « je ne sais pas ». Cette réponse vaut mieux qu'un signe inventé, une lacune comblée sans avertissement ou une traduction qui dissimule ses hésitations.

04 Pourquoi les hiéroglyphes résistent à la reconnaissance automatique

Le problème

Une inscription n'est pas une page imprimée avec des caractères régulièrement alignés. Le support peut être cassé, usé, remployé ou éclairé de biais. Un relief peut se confondre avec une fissure. Une inscription peut être gravée en profondeur ou en relief. Deux signes peuvent se toucher ; un seul signe peut être brisé en plusieurs morceaux. Le décor voisin peut ressembler à un hiéroglyphe sans appartenir au texte.

La difficulté continue après l'image. Les signes peuvent être groupés dans un cadrat, ce petit espace carré ou rectangulaire où plusieurs signes sont arrangés ensemble. Le texte peut aller de droite à gauche ou de gauche à droite, en lignes mais aussi en colonnes de haut en bas. Des inversions honorifiques placent parfois un nom divin en tête sans changer l'ordre grammatical attendu.

EXEMPLE Un oiseau regarde généralement vers le début de la ligne. Cet indice aide à trouver le sens de lecture, mais une scène symétrique peut contenir des figures orientées dans les deux sens. L'indice devient alors une proposition, pas une preuve définitive.

Enfin, reconnaître une forme ne suffit pas à comprendre sa fonction. Un même signe peut noter un son, préciser le sens d'un mot ou participer à une graphie particulière. C'est pourquoi **Loracraft** sépare ce qui est vu, ce qui est reconnu et ce qui est interprété.

05 Voir n'est pas encore lire

Notre méthode

Loracrafft range chaque information à son bon niveau. Une zone sombre ou un contour appartient à l'observation. « Ce contour pourrait être le signe G17 » appartient à l'hypothèse. « Ici, ce signe vaut *m* » appartient à l'analyse linguistique. Le mot choisi en français appartient à la traduction.

Cette séparation paraît méticuleuse, mais elle empêche une erreur précoce de devenir invisible. Si une identification change, le programme peut montrer quelles analyses et quelles phrases françaises en dépendent. L'image et les contours, eux, ne sont pas réécrits pour donner raison à la nouvelle lecture.

EN CLAIR Le programme ne remplace pas une ancienne proposition par une nouvelle comme si la première n'avait jamais existé. Il ajoute la nouvelle proposition et relie les deux. Le chercheur peut ainsi reconstituer l'histoire de l'analyse.

Résultat attendu

À la fin, on ne dispose pas seulement d'une traduction. On dispose d'un dossier où chaque segment traduit peut être relié aux signes, aux zones de l'image, aux validations et aux parallèles qui le soutiennent.

État actuel

Le principe est déjà inscrit dans *LoraDocument 1.0* : objets de mise en page, hypothèses de signes et validations humaines sont enregistrés séparément. La chaîne linguistique complète reste à construire.

06 Ce que devient une photographie lorsqu'elle entre dans Loracraftt

Le problème

Une photographie peut être renommée, recadrée ou compressée sans que l'on sache ensuite laquelle a servi à une analyse. Or une différence minuscule d'éclairage ou de cadrage peut modifier les contours détectés. Le fichier d'origine doit donc rester identifiable.

Notre méthode

À l'import, **Loracraftt** conserve l'original, relève ses dimensions et les informations disponibles, documente les droits, puis calcule une empreinte numérique. Cette empreinte est une sorte de signature digitale du fichier : si un seul octet change, l'empreinte change aussi. Plusieurs vues du même monument peuvent être reliées sans être confondues.

EXEMPLE Une vue d'ensemble situe l'inscription sur le monument ; une photographie en lumière rasante révèle des traits ; un détail à haute résolution aide à examiner un signe. Ce sont trois fichiers différents, rattachés au même dossier scientifique.

Résultat attendu

Un tiers doit pouvoir retrouver exactement la source utilisée, comprendre ses conditions d'accès et savoir quelles images de travail en dérivent.

État actuel

Le noyau sait déjà conserver la provenance des images dans le workflow local. Le module complet d'acquisition, avec catalogues, licences et contrôles systématiques, appartient encore à l'architecture cible.

07 L'original ne bouge jamais



test_5 : photographie réelle d'une stèle, utilisée pour éprouver le redressement automatique (Karnak).

Loracrafft travaille sur des copies. Une version peut être redressée, une autre contrastée, une troisième préparée pour un moteur particulier. Chacune garde le lien vers l'original et les paramètres qui l'ont produite.

POURQUOI EST-CE IMPORTANT ? Une image plus spectaculaire n'est pas forcément plus fidèle. Un contraste trop fort peut faire disparaître un trait fin ; un débruitage peut effacer une trace ancienne. Conserver les étapes permet de comparer et de revenir en arrière.

08 Redresser l'image avant de la découper

Le problème

Si le programme cherche des lignes et des colonnes sur une photographie inclinée, il fabrique une géométrie fausse. Les limites de registres deviennent diagonales, les cadrats se chevauchent et les vignettes de signes sont mal cadrées. La reconnaissance ne peut pas réparer un découpage incohérent.

Notre méthode

Le prototype repère d'abord des segments rectilignes visibles dans l'image. Il estime ensuite l'orientation dominante tout en tenant compte du fait qu'une inscription peut être organisée verticalement. En langage courant : il cherche l'inclinaison générale sans décider que tout texte doit être horizontal.

EXEMPLE Sur une première version de l'image test_5, l'inclinaison (erronée) mesurée était d'environ 44,7 degrés. Une première version avait lancé le découpage trop tôt. Après déplacement du redressement en amont, la stèle a retrouvé une orientation verticale exploitable.

Résultat attendu

Le programme fournit une image de travail redressée et note l'angle détecté, l'angle réellement appliqué et la transformation qui permet de revenir aux coordonnées de la photographie d'origine.

09 Redressement, perspective et contraste : trois opérations différentes

Redresser

On corrige une rotation : le monument est photographié de travers, mais son plan reste approximativement parallèle au capteur.

Corriger la perspective

On compense un effet de trapèze : un bord du monument est plus éloigné que l'autre. Cette opération exige de connaître ou de proposer les quatre coins de la surface utile.

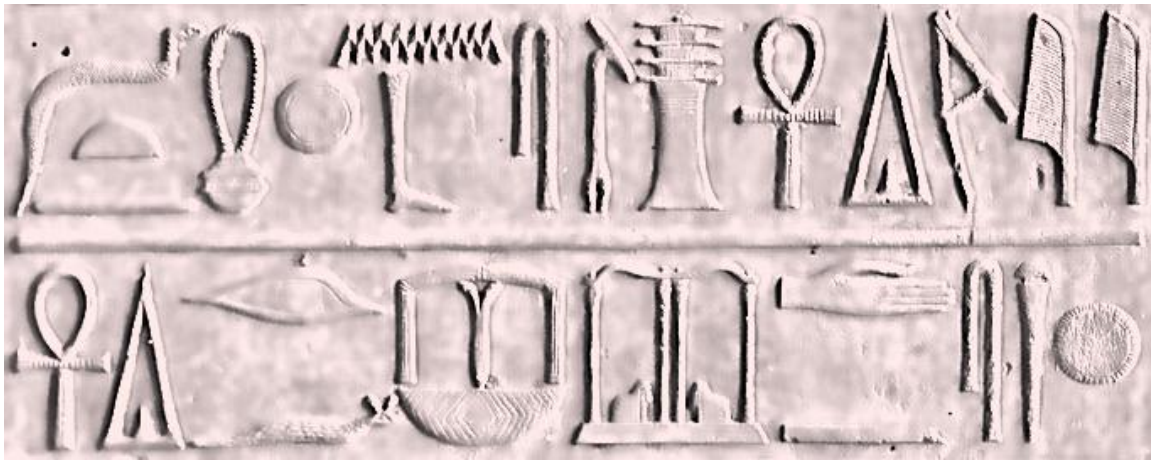
Améliorer la lisibilité

On ajuste localement le contraste ou on réduit certains bruits afin que les reliefs faibles soient plus visibles. Le résultat est une aide à l'analyse, jamais une nouvelle source.

LIMITE ACTUELLE Le redressement automatique est opérationnel sur le cas testé avec le prototype actuellement disponible. La détection entièrement automatique des quatre coins et la correction de perspective ne le sont pas encore. Un bord cassé, un décor ou un élément moderne peut tromper l'estimation ; une vérification visuelle reste indispensable pendant le développement.

Plusieurs préparations peuvent coexister. Un moteur de contours n'a pas forcément besoin de la même image qu'un lecteur humain. Le dossier conserve ces variantes sans les confondre.

10 Une même inscription, deux apparences



test_2 : version modifiée³ de l'image suivante pour faire ressortir les reliefs et les limites de structure.



*test_3 : version naturelle du même type d'inscription à deux registres
(Karnak, chapelle blanche de Sésostriis 1er).*

À RETENIR Une image préparée peut aider un algorithme sans être meilleure pour tous les usages. **Loracraft** conserve l'image source, les dérivés et la recette de transformation afin que l'on sache toujours ce que l'on regarde.

³ Toutes les modifications d'images ont été réalisées par le logiciel ChatGPT 5.6 d'OpenAI selon les instructions de l'auteur.

11 Repérer la structure avant les signes

Le problème

Une inscription possède une organisation. Elle peut comporter plusieurs registres superposés, des colonnes, des lignes, des cartouches et des zones endommagées ou manquantes. Si l'on découpe immédiatement chaque petite tache, on perd cette organisation et l'on obtient des fragments sans contexte.

Notre méthode

Loracrafft procède du grand vers le petit : surface inscrite, registres ou grandes bandes, lignes ou colonnes, cadrats, puis signes. Cette stratégie est dite « registres d'abord ». Elle a été adoptée après l'échec d'un découpage global sur une inscription à deux lignes.

EN CLAIR Avant de chercher les lettres d'une page, on repère ses colonnes, ses paragraphes et ses lignes. Pour les hiéroglyphes, l'ordre exact est différent, mais l'idée est la même : comprendre la mise en page évite de mélanger des voisins qui ne doivent pas l'être.

Résultat attendu

Le programme produit une carte géométrique : chaque région possède une forme, une position, un type proposé et des relations avec les régions voisines. À ce stade, aucun code Gardiner n'est encore attribué.

12 Pourquoi une tache n'est pas un signe

L'expérience qui a changé le plan

Une première méthode isolait les composantes connexes, c'est-à-dire les groupes de pixels qui se touchent. Elle supposait implicitement qu'un groupe correspondait à un signe. Les images réelles ont invalidé cette idée.

Un signe gravé peut être interrompu par l'usure ou une cassure et former plusieurs composantes. À l'inverse, deux signes voisins peuvent se toucher et n'en former qu'une. Une fusion expérimentale a ramené 173 composantes à 56 groupes, mais elle fusionnait encore certains voisins et laissait séparés certains fragments.

EXEMPLE Un animal appartenant au décor sur l'image peut être une forme parfaitement nette sans être un hiéroglyphe. Le forcer dans la liste de Gardiner produirait une réponse techniquement assurée mais scientifiquement absurde.

Notre méthode

Le système distingue désormais les signes probables, les groupes, les fragments, le bruit et les objets inconnus. Les motifs valides mais étrangers au texte doivent aussi être signalés. Ces catégories orientent la revue au lieu de forcer toute forme vers un reconnaisseur.

État actuel

Le prototype détecte correctement les deux registres de test_2, mais il ne délimite pas encore toutes les véritables unités hiéroglyphiques. Le découpage reste le principal verrou de la phase 1.

13 Le cas vertical : garder le contexte



*test_4 : inscription verticale et dégradée, utile pour éprouver colonnes, cartouches, fragments et ordre de lecture
(Pyramide d'Ounas - © Morris Franken).*

Une image verticale rappelle pourquoi il ne faut pas imposer les habitudes d'une page moderne. Les séparateurs de colonnes, les figures orientables et les cartouches donnent des indices qui doivent rester reliés à leur emplacement.

POURQUOI EST-CE IMPORTANT ? Une vignette isolée peut ressembler à plusieurs signes. Sa colonne, ses voisins et sa position dans un groupe réduisent parfois l'ambiguïté. Le contexte ne doit donc pas être jeté au moment du découpage.

14 Trouver le sens de lecture

Le problème

Comme on sait, les hiéroglyphes peuvent être disposés horizontalement, et se lire de droite à gauche ou de gauche à droite, ou verticalement. Dans un texte vertical, les signes se lisent de haut en bas, mais il faut encore déterminer dans quel ordre parcourir les colonnes.

Le sens de lecture est d'abord indiqué par les hiéroglyphes représentant des êtres animés : personnages, oiseaux et autres animaux regardent presque toujours⁴ vers le début du texte. Le lecteur va donc à leur rencontre. L'organisation des lignes, des colonnes, des cadrats et des séparateurs apporte des indications supplémentaires.

Une scène peut cependant contenir plusieurs inscriptions indépendantes. Un souverain et une divinité qui se font face peuvent, par exemple, être accompagnés de textes orientés en sens opposés.

Il ne faut donc pas chercher à imposer un sens unique à toute la représentation : chaque zone inscrite doit d'abord être délimitée et étudiée séparément.

Notre méthode

Loracrafft commence par repérer les différentes zones de texte présentes dans l'image. Pour chacune d'elles, il examine l'orientation des signes animés, la disposition en lignes ou en colonnes et l'organisation des signes à l'intérieur des cadrats.

Le logiciel propose ensuite un ordre de lecture pour chaque inscription.

Si une partie est effacée, endommagée ou difficile à distinguer, il peut conserver plusieurs hypothèses. Leur cohérence est vérifiée à l'aide du vocabulaire, de la grammaire et de textes comparables provenant d'autres monuments.

⁴ Il existe toutefois de rares inscriptions dites *rétrogrades*, dans lesquelles les signes regardent vers la fin du texte. L'organisation des lignes, des colonnes et des cadrats, ainsi que la grammaire, permettent alors de vérifier ou de corriger la première hypothèse (cf. Henry George Fisher, <https://www.metmuseum.org/fr/met-publications/the-orientation-of-hieroglyphs-part-1-reversals>).

EXEMPLE Dans une scène où le pharaon et un dieu se font face, les inscriptions placées devant chacun d'eux peuvent se lire en sens opposés. L'orientation des personnages ne suffit pas à elle seule. **Loracraftt** observe surtout les signes composant chaque texte, puis traite séparément les paroles du dieu, les titres du souverain et la formule décrivant l'offrande.

15 Proposer un code Gardiner sans inventer

Le problème

Les égyptologues emploient couramment les codes Gardiner comme G17 ou N35 pour désigner des signes d'une liste de référence organisée par grandes catégories. Ce code est pratique, mais une forme usée ou mal découpée peut correspondre à plusieurs entrées.

Notre méthode

Pour chaque objet, *LoraGlyph* peut produire une courte liste de candidats. Chaque candidat indique le référentiel consulté, le moteur utilisé et le type de score. Si aucun candidat sérieux n'existe, la liste reste vide et l'objet part en revue.

EN CLAIR Le programme ne dit pas : « ce signe est G17 ». Il dit : « pour cette vignette et avec ce catalogue, G17 est une possibilité classée devant deux autres avec tel score ». La décision scientifique vient ensuite.

Résultat attendu

Le logiciel voit la vignette, la région dans l'image complète, les candidats et les raisons du classement. Il peut accepter, rejeter, corriger, fusionner des fragments ou demander un nouveau découpage.

État actuel

LoraGlyph 1.0.1 produit déjà des vignettes et des hypothèses séparées des objets géométriques. La qualité de reconnaissance n'est pas encore validée scientifiquement.

16 Comparer des formes : utile, mais fragile

Les premières expériences

Le prototype compare des silhouettes, des contours, des répartitions de pixels et sept nombres qui résument la forme générale. Ces derniers sont souvent appelés *moments de Hu*⁵. Ils aident à retrouver une forme malgré un changement de taille ou de position.

Aucune de ces mesures ne suffit seule. Un contour mal découpé donne une mauvaise comparaison ; deux signes différents peuvent posséder une silhouette proche ; une variation d'épaisseur ou de cadrage peut faire baisser le score d'un exemple évident.

EXEMPLE Un essai sur le signe G1 a donné de façon erronée une reconnaissance de 0,748 % avec une comparaison brute, alors que le cas paraissait visuellement simple. Ce résultat interdit de décréter qu'un seuil unique, par exemple 0,75, sépare partout le vrai du faux.

La suite prévue

Plusieurs méthodes seront évaluées séparément, puis éventuellement combinées. Le développement d'un réseau de neurones n'est pas prioritaire tant que les images de référence ne sont pas correctement découpées et que le corpus n'est pas organisé en ensembles fiables pour l'apprentissage et l'évaluation. Entraîné sur de mauvaises vignettes, un tel modèle apprendrait principalement à reproduire nos erreurs.

PRINCIPE La puissance du moteur ne compense ni une mauvaise géométrie ni un corpus mal construit.

⁵ Ming-Kuei Hu (1918–2002) est un ingénieur et professeur sino-américain de l'université de Syracuse. Il publia en 1962 une méthode décrivant les formes au moyen de sept valeurs mathématiques, aujourd'hui appelées *moments invariants de Hu*. Ces valeurs permettent de comparer des silhouettes indépendamment de leur position, de leur taille et, dans une large mesure, de leur orientation.

17 Les 6 722 fichiers de glyphes : une bibliothèque, pas 6 722 signes.

La ressource disponible

Une base locale issue des ressources de Raymond Monfort⁶ contient 6 722 fichiers glyphiques. Ce chiffre compte des fichiers et des variantes graphiques. Il ne signifie pas que la liste de Gardiner possède 6 722 signes distincts.

Le travail nécessaire

Chaque fichier doit être inventorié : code associé, variante, format, dimensions, empreinte, source, licence et éventuels doublons. Les formes libres de droits, les références consultables et les ressources non redistribuables doivent être distinguées.

POURQUOI EST-CE IMPORTANT ? Un catalogue volumineux peut donner une impression de couverture alors qu'il contient beaucoup de variantes d'un même signe, des fichiers mal nommés ou des droits incertains. La quantité ne remplace pas la qualification.

État actuel

Les fichiers sont disponibles pour l'expérimentation, mais le manifeste scientifique et juridique reste à produire. En parallèle, un lot de 160 scans réels peut servir à l'évaluation, à condition de garder toutes les vues d'une même inscription dans la même partie du jeu de données.

⁶ <https://papocle.com/RaymondMonfort/>

18 L'ordinateur propose, le chercheur décide

Le problème

Une correction humaine ne doit pas effacer l'hypothèse automatique ni devenir automatiquement une vérité universelle. Deux spécialistes peuvent examiner la même région et ne pas être d'accord. Leur désaccord contient une information scientifique utile.

Notre méthode

Chaque action de revue devient un événement daté : accepter, rejeter, corriger, déclarer illisible, fusionner, scinder ou reporter. L'action précise sa cible, son auteur, son degré de certitude et, si possible, sa justification.

EXEMPLE Un moteur propose G43. Un premier annotateur accepte avec réserve ; un second propose G17 et justifie son choix par le contexte. **Loracraft** conserve les deux avis. Une décision éditoriale peut retenir G17 pour une version publiée sans effacer G43 ni le premier avis.

Résultat attendu

La revue est une contribution historisée. On peut calculer un accord entre annotateurs, mais le nombre de voix ne remplace pas l'argument.

État actuel

Le modèle de validations multiples existe dans *LoraDocument 1.0*. Une ancienne revue en console écrit déjà un journal de corrections. L'interface visuelle permettant de redessiner, fusionner et scinder les régions reste prioritaire.

19 Ce que devra montrer l'écran de revue

Image complète. Elle conserve le contexte monumental et la position du signe.

Région proposée. Sa vignette reste visible et son contour peut être corrigé.

Images de travail. Elles aident l'examen sans masquer la photographie originale.

Candidats Gardiner. L'écran montre leur ordre, leurs scores et le moteur qui les a produits.

Décisions antérieures. Les désaccords et les justifications restent consultables.

Actions de revue. Accepter, rejeter, corriger, fusionner, scinder, déclarer illisible ou reporter.

EN CLAIR La priorité n'est pas un bouton « traduire ». C'est un écran où l'on peut corriger proprement le découpage de test_2 (par exemple) et enregistrer cette correction sans écraser le passé.

Une validation ne rejoint pas automatiquement les données d'entraînement. Elle doit d'abord être contrôlée, assortie de droits compatibles et placée dans une partition qui ne contamine pas les tests.

20 Du signe validé au Manuel de Codage

Le problème

Une liste linéaire de codes ne suffit pas toujours à représenter une inscription. Dans un cadrat, des signes peuvent être placés au-dessus, au-dessous ou côte à côte. Il faut conserver cette organisation et la relier aux objets visibles.

Notre méthode

Le Manuel de Codage, abrégé MdC, fournit des conventions pour noter des signes et leurs groupements. **Loracrafft** construit une représentation structurée : chaque élément de l'encodage restera aligné sur le ou les objets visuels dont il provient.

EXEMPLE Si deux signes sont superposés dans un cadrat, l'encodage doit indiquer leur relation. Si une partie est lacunaire, la lacune reste marquée. Si deux ordres sont possibles, deux branches peuvent coexister.

Résultat attendu

Le chercheur peut afficher côte à côte l'image, les cadrats, l'encodage MdC et un rendu hiéroglyphique. Une correction de signe modifie seulement les représentations qui en dépendent.

État actuel

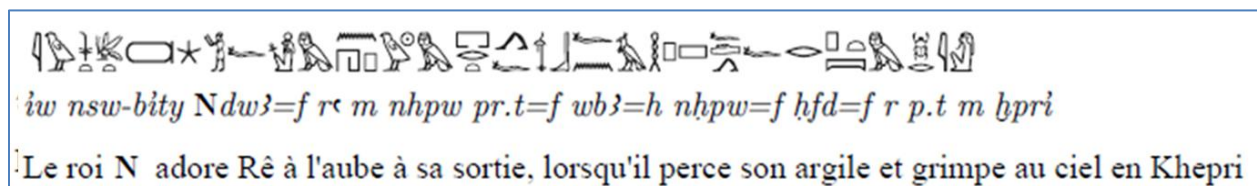
LoraMdC appartient encore à l'architecture cible. Son développement vient après la stabilisation de la segmentation, de l'ordre de lecture et des validations de signes.

21 Translittérer n'est pas traduire

La translittération

Elle représente en caractères latins spécialisés les valeurs retenues pour les signes. Elle facilite l'analyse linguistique, mais elle est déjà le résultat de choix : valeur phonétique, complément, déterminatif, ordre du groupe et restitution éventuelle.

Exemple :



(Livre des Portes, sarcophage de Séthi Ier, XIXe dynastie ; transcription, translittération et traduction de Raymond Monfort, <http://raymond.monfort.free.fr>).

Le découpage en mots

Les inscriptions ne marquent pas les mots comme le français moderne. Le programme doit proposer où un mot commence et où il finit. Deux segmentations peuvent conduire à deux analyses différentes.

L'analyse

Pour chaque mot proposé, **Loracrafft** cherchera un lemme (la forme de dictionnaire), une catégorie grammaticale, des traits morphologiques et des relations syntaxiques. Chaque proposition renverra à une règle ou à une ressource identifiable.

EN CLAIR Passer des codes Gardiner au français ne se fait pas en un saut. Entre les deux se trouvent l'ordre, le groupement, les valeurs, la translittération, les mots, la traduction, la grammaire et le contexte.

État actuel

Le module de grammaire n'est pas encore implémenté. Le module du lexique est en cours de validation (base SQL de 7830 entrées conçue à partir du dictionnaire de Yvonne Bonnamy version 2019). Le document technique décrit leurs contrats afin qu'ils puissent être construits sans confondre dictionnaire, hypothèse et décision éditoriale.

22 Faire voyager l'incertitude jusqu'au français

Le problème

Un signe incertain peut modifier un mot ; ce mot peut modifier la grammaire ; la grammaire⁷ peut modifier la traduction. Si l'incertitude disparaît au milieu de la chaîne, la phrase française semble plus sûre que les données ne le permettent.

Notre méthode

Chaque résultat déclare les éléments dont il dépend. Lorsque deux lectures restent plausibles, le programme peut produire deux branches. Une synthèse compare ces branches et indique précisément ce qui change.

EXEMPLE La lecture A produit une construction bien attestée. La lecture B exige une graphie rare mais reste compatible avec la zone abîmée. L'édition peut préférer A, conserver B en variante et expliquer pourquoi.

Lacunes et restitutions

Une zone détruite est une observation matérielle. Le texte proposé pour la combler est une restitution. Les deux ne doivent jamais être enregistrés comme une seule chose. Dans l'export, le texte restitué reste marqué et relié à ses arguments.

POURQUOI EST-CE IMPORTANT ? La prudence n'interdit pas de conclure. Elle oblige seulement à montrer où finit l'observation et où commence la reconstruction.

⁷ Le prototype actuel utilise un mix Champollion 1836 - Gardiner 1957 - Malaise-Winand 1999, avec une (grosse) pincée de Grandet-Mathieu.

23 Comment le TLA peut aider

Ce qu'est le TLA dans notre chaîne

Le **Thesaurus Linguae Aegyptiae** est le plus vaste corpus électronique mondial de textes égyptiens lemmatisés ; il permet de consulter les textes, les lemmes, leurs attestations et leurs contextes. Il n'est pas un service auquel on envoie une photographie pour qu'il reconnaisse directement les signes. Il intervient après les premières hypothèses visuelles et linguistiques.

Notre méthode

Loracrafft prépare une requête à partir de translittérations possibles, de séquences de signes, de mots, de périodes et de types d'objets. Un connecteur interrogera uniquement les services ou données officiellement accessibles et conservera la version de la ressource, la date de consultation et les identifiants externes.

EXEMPLE Une séquence incertaine pourrait correspondre à une formule de stèle du Moyen Empire. La recherche retrouve plusieurs attestations. Ces passages deviennent des éléments de comparaison ; ils ne transforment pas automatiquement l'hypothèse en fait.

Résultat attendu

Chaque rapprochement indique ce qui a été comparé, pourquoi il est classé à cette place et où consulter le passage cité, dans la limite des droits.

24 Un parallèle est un indice, pas un tampon de vérité

Le risque

Trouver une formule très proche peut renforcer une lecture, mais l'absence de résultat ne prouve pas que la lecture est impossible. Le corpus peut être incomplet, la graphie différente ou la requête trop étroite.

Il existe aussi un risque de raisonnement circulaire. Une édition moderne peut avoir déjà utilisé l'inscription étudiée pour reconstruire un passage. Présenter cette édition comme une preuve indépendante reviendrait à faire confirmer le texte par lui-même.

Notre méthode

Le système distingue œuvre, témoin matériel, passage et édition. Il conserve la généalogie des sources, la version des index et les règles de normalisation. Les résultats anciens ne sont pas réécrits lorsqu'un corpus évolue ; ils sont signalés comme calculés avec une version antérieure.

LIMITE ACTUELLE *LoraCorpus* et le connecteur TLA restent à construire. Avant l'indexation à grande échelle, il faut fixer les droits, les méthodes de normalisation et une politique d'évaluation par des spécialistes.

Résultat attendu

Le chercheur reçoit une liste de passages comparables, pas une sentence. Il peut vérifier la source, écarter un rapprochement et expliquer pourquoi un autre est pertinent.

25 Mettre les arguments sur la table

Le problème

À la fin de la chaîne, plusieurs indices peuvent converger ou se contredire : forme du signe, ordre de lecture, grammaire, lexique, datation et passages parallèles. Une note globale cacherait la nature de ces arguments.

Notre méthode

LoraReasoner construira des scénarios. Chaque conclusion indiquera ses prémisses, les règles appliquées, les contre-arguments et les hypothèses nécessaires. Les règles géométriques, linguistiques, documentaires et statistiques resteront identifiables.

EXEMPLE Lecture A préférée parce qu'elle concorde avec la forme, la syntaxe et deux parallèles ; lecture B conservée parce que la zone est endommagée ; lecture C rejetée parce qu'elle impose un ordre impossible.

Résultat attendu

Le système peut répondre à une question contrefactuelle : si G43 était remplacé par G17, quelles analyses et quelles phrases changeraient ?

État actuel

Ce module appartient à la cible de long terme. Une première version fondée sur des règles explicites et un graphe de dépendances est préférable à une boîte noire plus impressionnante mais moins contrôlable.

26 Produire plusieurs traductions utiles

Trois usages

Une traduction littérale montre la structure et les répétitions. Une traduction philologique explique les choix, les lacunes et les termes techniques. Une traduction fluide vise une lecture française naturelle⁸. Elles répondent à des besoins différents et peuvent coexister.

Notre méthode

Chaque segment français reste aligné sur la translittération et l'analyse qui le soutiennent. Les restitutions et les alternatives ne disparaissent pas. Un modèle génératif pourra aider à reformuler ou à comparer des solutions, mais il ne pourra ni inventer une source ni masquer une incertitude.

POURQUOI EST-CE IMPORTANT ? Une phrase élégante peut être fausse avec beaucoup d'assurance. **Loracraft** cherche d'abord une traduction vérifiable, puis une formulation lisible.

Validation humaine

La version publiée est revue par un spécialiste pendant la phase de développement. Le système conserve les propositions automatiques et les corrections, afin que l'on puisse comprendre une révision ultérieure.

État actuel

LoraTranslate n'est pas encore implémenté. Après la reconnaissance des signes, leur mise en ordre, la translittération et l'analyse linguistique, il produira une traduction en français, en anglais et en allemand.

⁸ Nous avons choisi de passer du translittéré au français, du fait de la disponibilité plus importante des données nécessaires à la base de connaissance en cette langue (lexique et grammaire notamment), puis le logiciel proposera une traduction en anglais ou en allemand selon les réglages de l'utilisateur.

27 Corriger tard sans tout recommencer

Le scénario

Une édition de travail a retenu G43. Après examen, un chercheur propose G17. Cette correction affecte la translittération, deux mots possibles, une analyse grammaticale et une phrase française. Elle n'affecte ni la photographie, ni le redressement, ni les autres colonnes.

Notre méthode

Le programme crée une nouvelle validation et une nouvelle branche de travail. Il recalcule seulement les résultats qui dépendent du signe corrigé. Les étapes indépendantes restent valides. Une nouvelle révision est produite ; l'ancienne reste consultable et citable.

EN CLAIR On ne repeint pas tout le chantier parce qu'une pierre a été mieux identifiée. On remonte la chaîne des conséquences et l'on refait uniquement ce qui doit l'être.

Résultat attendu

Une comparaison de révisions indique le signe modifié, la translittération affectée, l'analyse devenue préférable et la traduction mise à jour. Le lecteur peut vérifier la justification du changement.

État actuel

Le prototype sait déjà enregistrer chaque nouveau résultat sans écraser les précédents. Lorsqu'un utilisateur corrigera une zone, un signe ou un mot, **Loracraft** devra ensuite recalculer uniquement les résultats qui dépendent de cette correction, jusqu'à la traduction, sans reprendre inutilement tout le traitement depuis le début.

28 Les cinq images de test comme carnet de bord

Les images de test n'ont pas toutes le même rôle. La ligne propre vérifie les bases. L'inscription à deux registres révèle les erreurs de segmentation. Sa variante visuelle aide à comparer les préparations. Le document vertical éprouve colonnes, cartouches et dégradation. La stèle photographiée de travers teste le redressement sur un objet réel.

Cas	Question posée	Règle de prudence
test_1	Conserve-t-on les détections déjà justes ?	Aucune amélioration ne doit créer une régression.
test_2	Les deux registres et chaque unité sont-ils correctement délimités ?	Ne pas traiter test_3 tant que ce niveau n'est pas validé.
test_3	La méthode tient-elle sur une variante plus naturelle ?	Cas volontairement différé.
test_4	La structure verticale et dégradée reste-t-elle lisible ?	Garder le contexte et les alternatives.
test_5	Le redressement précède-t-il bien la géométrie ?	Toujours contrôler visuellement l'orientation.

29 Où en est réellement le prototype

Niveau	Ce que cela signifie aujourd'hui
Déjà tangible	<i>LoraDocument 1.0</i> ; redressement de l'image et amélioration du contraste ; séparation des deux registres de l'image test_2 ; production de vignettes et de plusieurs hypothèses de reconnaissance ; possibilité de valider plusieurs résultats ; enregistrement local et versionné des résultats sans écraser les précédents.
Expérimental	Découpage des cadrats et signes ; cartouches ; comparaison par gabarits et descripteurs de forme ; classement Gardiner.
À construire	Ordre de lecture complet ; MdC ; grammaire ; lexique ; recherche de corpus et connecteur TLA ; raisonnement ; traduction.
Option ultérieure	Restitution sonore, clairement présentée comme conventionnelle ou reconstruite.

30 Ce qui fonctionne, ce qui échoue, ce que nous avons appris

Ce qui fonctionne

Le noyau *LoraDocument 1.0*, l'écriture sans écrasement, les identifiants stables, la provenance des images, les validations multiples et une orchestration locale sous Windows 11 et Python 3.12. Le redressement de test_5 a validé l'ordre général de la préparation visuelle.

Ce qui échoue encore

Le découpage exact des cadrats et des signes sur test_2. Les cartouches ouverts ou coupés par le bord restent difficiles. Les comparaisons de forme sont fragiles lorsque le cadrage, le bruit ou l'épaisseur changent. Aucun score actuel ne doit être présenté comme une certitude scientifique.

Ce que l'échec a appris

La structure doit être détectée avant les petites unités ; une composante de pixels n'est pas un signe ; les objets étrangers au texte doivent pouvoir rester étrangers ; le rapport d'exécution et les sorties doivent être archivés par version et par cas de test.

Ce qui n'est pas commencé

L'ordre de lecture complet, le MdC, la grammaire, le lexique, la recherche de parallèles, le raisonnement et la traduction sont des modules projetés. Les décrire précisément dans l'**Architecture scientifique** ne signifie pas qu'ils sont implémentés.

31 La feuille de route en quatre phases

1. Voir et valider : stabiliser la géométrie de test_2, construire l'interface de revue et mesurer les résultats sur des inscriptions réelles.
2. Encoder et analyser : produire le MdC aligné, la translittération, les mots et des analyses grammaticales sourcées.
3. Comparer et argumenter : rechercher des parallèles reproductibles, notamment via les ressources autorisées du TLA, puis construire des scénarios explicables.
4. Traduire et publier : proposer plusieurs niveaux de traduction, les soumettre à une revue humaine et publier une révision entièrement traçable.

PRIORITÉ IMMÉDIATE Terminer la vérité terrain de test_2 et l'interface de correction géométrique avant de chercher un reconnaisseur plus spectaculaire, puis passer au test_3.

Les fondations (dossier commun, provenance, droits, tests, archivage et migrations) accompagnent toutes les phases. Elles ne sont pas une phase décorative : elles rendent les résultats réutilisables.

32 À quoi reconnâitrons-nous que Loracraftt est utile ?

Loracraftt sera utile lorsque l'outil d'un chercheur pourra partir d'une photographie, corriger la géométrie, examiner les propositions de signes, construire une lecture, retrouver des comparaisons et publier une traduction sans perdre la trace de ses choix.

Son ambition n'est pas de supprimer l'incertitude propre aux inscriptions anciennes. Elle est de la rendre visible, localisée et productive. Une alternative conservée aujourd'hui peut devenir la meilleure lecture demain, grâce à une photographie nouvelle, un remontage de fragments ou un parallèle mieux compris.

L'IDÉE CENTRALE Une image, une région, un signe proposé, une analyse et une traduction ne sont pas cinq versions successives de la même vérité. Ce sont cinq objets différents reliés par des preuves.

Le logiciel pourra changer de moteur, de catalogue ou d'interface. Ce qui doit durer est la mémoire scientifique : ce qui a été observé, ce qui a été proposé, qui a décidé, sur quelles preuves et dans quelle version.

Le prochain succès n'est donc pas une traduction automatique de démonstration. C'est une segmentation correcte et révisable de test_2, examinée par des humains et conservée dans un dossier que l'on pourra rejouer. À partir de ce socle, chaque couche supplémentaire aura un sens.

33 Petit glossaire

Terme	Définition courante
Cadrat	Petit espace graphique où un ou plusieurs signes sont organisés ensemble.
Code Gardiner	Code de référence utilisé pour désigner un signe dans une liste classée par catégories.
Corpus	Ensemble organisé de textes et de témoins utilisé pour chercher des attestations et des parallèles.
Empreinte numérique	Valeur calculée à partir d'un fichier qui permet de vérifier son identité et de détecter une modification.
Hypothèse	Proposition révisable, par exemple l'identification possible d'un signe.
Lemme	Forme de dictionnaire à laquelle une forme rencontrée dans le texte est rattachée.
Manuel de Codage (MdC)	Convention pour représenter les signes hiéroglyphiques et leur disposition.
Parallèle	Passage comparable qui peut éclairer une lecture sans la valider automatiquement.
Provenance	Histoire documentée d'un fichier, d'une transformation, d'une proposition ou d'une décision.
Registre	Grande bande horizontale ou zone structurée d'une composition.
Translittération	Représentation des valeurs linguistiques égyptiennes avec un alphabet spécialisé.
Vérité terrain	Découpage ou annotation de référence établi et contrôlé par des humains pour évaluer le programme.

34 Sources, limites et articulation avec le volume technique

Source principale

Architecture scientifique de Loracrafft, Didier DTL Morandi, version 1.1 du 17 août 2026.

Le présent volume en reformule le contenu pour un lecteur non spécialiste. En cas de besoin d’implémentation, le volume technique demeure la référence des contrats, formats, interfaces et règles de nommage.

Ressources mentionnées

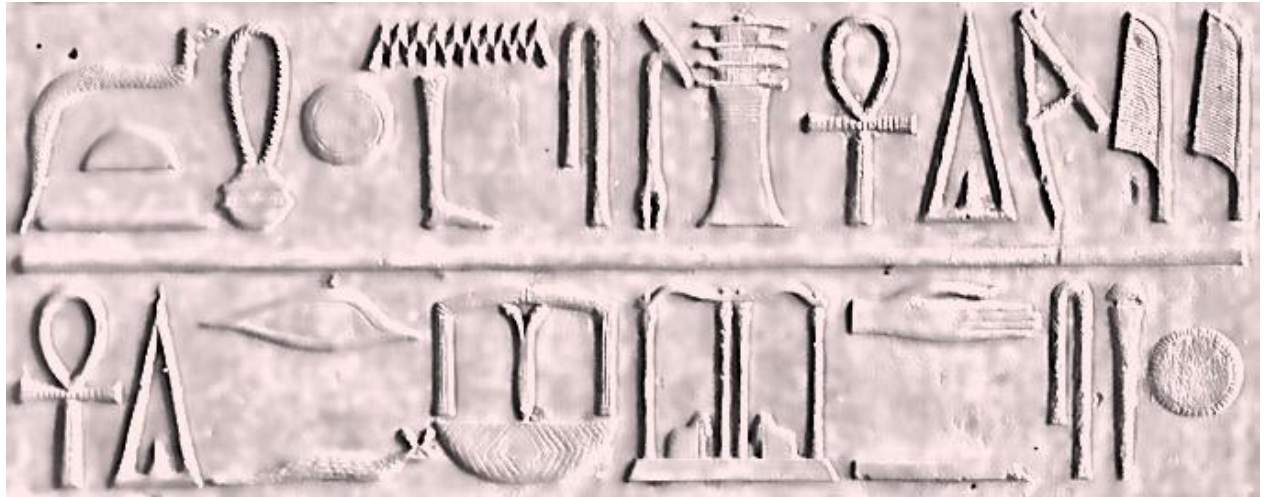
Liste de Gardiner pour la classification des signes ; Manuel de Codage pour leur représentation ; *Thesaurus Linguae Aegyptiae* pour les lemmes, attestations et corpus accessibles ; ressources glyphiques de Raymond Monfort ; standards d’édition, d’image et de provenance cités dans le volume technique.

Limites de ce document

Les résultats chiffrés décrivent des expériences de développement conduites à ce jour. Ils ne constituent pas encore des mesures généralisables à un corpus représentatif. Les cinq images sont des cas de travail, non une preuve de performance globale.

Règle pour les notes futures

Les notes de bas de page doivent servir aux sources, aux précisions historiques et aux discussions scientifiques. Une définition nécessaire à la compréhension reste dans le texte. Un contrat de données, un exemple JSON ou une interface détaillée retourne dans le volume technique.





test_5.png

