Le projet DHARMA et le renouvellement de l'épigraphie des mondes indien et indianisé

Arlo Griffiths

▶ **To cite this version:**

Arlo Griffiths. Le projet DHARMA et le renouvellement de l'épigraphie des mondes indien et indianisé. L'École française d'Extrême-Orient : 120 ans de recherches en Asie. Actes du colloque organisé par l'École française d'Extrême-Orient, sous le haut patronage de l'Académie des Inscriptions et Belles-Lettres à l'Académie des Inscriptions et Belles-Lettres, le 3 décembre 2021, Académie des inscriptions et belles-lettres, pp.71-86, 2023. hal-04135524

HAL Id: hal-04135524
https://hal.science/hal-04135524v1

Submitted on 22 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'École française d'Extrême-Orient : 120 ans de recherches en Asie

Christophe Marquet et Nicolas Grimal éd.

ACADÉMIE DES INSCRIPTIONS ET BELLES-LETTRES

ACTES DE COLLOQUE

L'École française d'Extrême-Orient : 120 ans de recherches en Asie, p. 71-86.

LE PROJET DHARMA ET LE RENOUVELLEMENT DE L'ÉPIGRAPHIE DES MONDES INDIEN ET INDIANISÉ

C'est un rare honneur que de pouvoir exposer devant cette académie consacrée aux inscriptions et belles-lettres l'état des lieux des recherches épigraphiques au sein de notre École, dont les premiers et plus prestigieux directeurs furent des spécialistes de l'étude des inscriptions de l'Asie du Sud-Est péninsulaire ! Suivant les pas des pionniers – Auguste Barth, Étienne Aymonier et Abel Bergaigne – qui avaient inauguré avant la fondation de notre École l'étude des inscriptions du Campā et du Cambodge, ce sont Louis Finot et George Cœdès qui ont consolidé et développé l'épigraphie de ces pays au travers de leurs œuvres titanesques, publiées pour la plus grande partie dans notre *Bulletin* ou dans nos collections de monographies[1]. Pour des raisons évidentes liées à l'histoire coloniale, l'épigraphie de l'Asie du Sud-Est insulaire n'était pas au cœur des recherches de notre École à leur époque mais nous devons néanmoins à Cœdès les remarquables articles parus sous les titres « Le royaume de Çrīvijaya » et « Les inscriptions malaises de Çrīvijaya », respectivement dans les volumes 18 et 30 de notre *Bulletin*[2]. Le savant y arrache aux ténèbres de l'oubli ce royaume bouddhique sis sur la route maritime entre l'Inde et la Chine et établit l'interprétation quasi définitive des plus anciens documents en langue malaise. Inutile d'insister ici sur tout ce que Cœdès a fait pour faire connaître l'histoire et la langue du Cambodge ancien.

Le travail épigraphique à leur époque commença par l'estampage des pierres inscrites sur le terrain, un art confié le plus souvent à des collaborateurs

1. Il suffit d'évoquer ici les vingt « Notes d'épigraphie » que publia Finot dans notre *Bulletin* entre 1902 et 1915, pour un total d'environ 400 pages ; les huit volumes de textes épigraphiques recueillis par Cœdès sous le titre *Inscriptions du Cambodge*, entre 1937 et 1966 ; ou encore les six portfolios de reproductions d'estampages, recueillis par les deux savants sous le titre *Inscriptions du Cambodge publiées sous les auspices de l'Académie des inscriptions et belles-lettres*, entre 1926 et 1937.

2. G. Cœdès, 1918 et 1930.

Fig. 1. – Estampeur inconnu, Phnom Penh, Cambodge (cliché EFEO CAM01705, M. Moride, 1952).

locaux (fig. 1) ; les estampages furent reçus à Paris ou à Hanoi, et permirent aux chercheurs français l'établissement des inventaires des inscriptions du Campā et du Cambodge[3]. Inventoriées, les inscriptions durent être déchiffrées. Rédigées soit en sanskrit, soit en langue vernaculaire – vieux cam, vieux khmer –, leur traduction posa ensuite un défi de taille variable, mais rarement négligeable. Le sanskrit était certes une langue relativement bien maîtrisée, mais il arrive souvent que les inscriptions rédigées en cette langue empruntent une forme de poésie savante dont l'interprétation est volontairement difficile. Les langues vernaculaires, elles, se révélèrent sensiblement différentes des parlers modernes, eux-mêmes alors encore peu connus ; toute aide à l'étude du vieux cam et du vieux khmer (dictionnaire, grammaire) faisait encore défaut.

3. G. Cœdès, 1908, 1923 et 1966.

Fig. 2. – N. Ramaswamy (« Babu ») du centre EFEO de Pondichéry réalise des photos d'une inscription à Tenneri, Tamilnadu, Inde (cliché N. Cane, 2012).

Avant de reprendre le fil des recherches dans le Sud-Est asiatique, arrêtons-nous brièvement sur les débuts de l'implication de notre École en épigraphie indienne. C'est naturellement une affaire plus récente qui commence tout aussi naturellement dans le Pays Tamoul. Le corpus épigraphique tamoul est le plus important de l'Inde ; il se compte en dizaines de milliers d'inscriptions, dont de très nombreuses inédites. C'est depuis les dernières décennies du XX^{e} siècle que des chercheurs de notre École, en poste à Pondichéry, se sont engagés dans la documentation d'inscriptions, le plus souvent gravées sur les murs des temples (fig. 2), et que l'épigraphie tamoule a commencé à jouer un rôle de plus en plus important dans leurs publications[4]. Ici comme dans d'autres domaines de notre activité scientifique en Inde, les chercheurs français collaborent intensivement avec

4. Je me borne ici à évoquer les deux tomes consacrés aux inscriptions du temple de Tiruvannamalai (P. R. Srinivasan et M.-L. Reiniche, 1990) ainsi que le recueil en deux volumes d'inscriptions pondichériennes (B. S. Kuppusamy, 2006 et 2010).

leur pairs Indiens (fig. 3), dont certains sont embauchés par notre École[5], ce qui contraste avec la situation en Asie du Sud-Est.

Les débuts de notre engagement en épigraphie tamoule correspondent, par coïncidence, avec la période où l'École a pu reprendre ses opérations au Cambodge et au Vietnam, après les décennies d'interruption causées par les guerres. Une fois l'accès aux pays rétabli, c'est grâce à l'initiative de Gerdi Gerschheimer (fig. 4), à l'époque directeur des études, que fut lancé le programme de « Corpus des inscriptions khmères », abrégé sous la forme de CIK. Permettez-moi de citer les mots qu'il consacra au programme dans le tome 90-91 du *Bulletin*[6] :

> « Le programme vise un double objectif. Il s'agit d'abord et avant tout **de reprendre et de poursuivre l'inventaire des inscriptions**, sur le modèle établi par G. Cœdès [...]. Depuis le dernier inventaire publié, qui s'arrêtait au numéro 1050 [...], de nombreuses épigraphes (**environ 200 à 250**) ont été découvertes, au Cambodge mais aussi en Thaïlande, au Laos ou au Vietnam – il en est même en Europe, aux USA ou au Japon. Elles seront progressivement intégrées **à l'inventaire général, désormais informatisé et couplé à l'inventaire des sites archéologiques** du Cambodge mené en collaboration entre l'EFEO et le ministère de la Culture du Cambodge. Mais cette opération ne saurait constituer un but en soi, et elle vise à faciliter et à **stimuler les recherches sur ce corpus**. Le programme entend donc également **inciter à la publication des inscriptions inédites, à la révision des éditions, traductions et annotations des inscriptions déjà éditées, enfin à l'exploitation de ce corpus pour des études thématiques et pluridisciplinaires**. Il s'agit en somme de forger un instrument qui, muni de toutes **les garanties de la rigueur philologique**, bénéficie des **apports de plusieurs disciplines** et soit en retour un **outil pour des recherches dans divers domaines** : linguistique (pour le vieux khmer, notamment), histoire, histoire des religions, histoire des sciences, architecture, archéologie, etc.
>
> Pour ce faire, un **corpus électronique des inscriptions** est en cours de constitution. Chaque épigraphe fait l'objet d'une notice descriptive aussi complète que possible (avec en particulier la bibliographie afférente, commentée, l'histoire des pérégrinations de l'inscription – quand on peut la restituer) ; elle comportera le texte et la traduction annotée de l'inscription. Ce corpus sera complété par un ensemble de **photographies d'estampages** (**en cours de numérisation** à l'EFEO) ou d'inscriptions, permettant aux chercheurs tout à la fois de vérifier l'exactitude des éditions, d'approfondir l'étude des inscriptions

5. Un bel exemple est la collaboration de Charlotte Schmid avec le Dr G. Vijayavenugopal, couronné dans le tome 92 de notre *Bulletin* par l'article « Au seuil du monde divin : reflets et passages du dieu d'Ālanturai à Puḷḷamaṅkai ».

6. G. Gerschheimer, 2003-2004.

FIG. 3. – Charlotte Schmid (EFEO) et G. Vijayavenugopal (EFEO Pondichéry)
lors de la Classical Tamil Summer School au centre EFEO de Pondichéry (cliché Ravi, 2011).

FIG. 4. – Gerdi Gerschheimer (EFEO/EPHE) dans
un dépôt archéologique à Phimai, Thaïlande (cliché P. Pichard, 2003).

et d'engager de nouvelles recherches (en particulier paléographiques). Enfin des fiches thématiques synthétiques (accessibles à partir des notices individuelles d'inscriptions), dans lesquelles seront données les annotations récurrentes sur des thèmes aussi divers que l'astronomie, la métrique, la mythologie, les mots à valeur numérique, permettront d'expliciter de manière plus approfondie qu'à l'ordinaire les présupposés nécessaires à l'intelligence des textes. **Ce corpus électronique pourra être constamment mis à jour et sera progressivement accessible, à moyen terme, sur le site internet de l'EFEO**, où figurera aussi, très prochainement, une Chronique des inscriptions du pays khmer recensant les publications nouvelles, les conférences, colloques et cours, les découvertes. »

Nombreux sont ceux parmi nous dont les recherches, qu'elles soient consacrées au pays Khmer ou non, ont été orientées, de façon plus ou moins directe, par l'agenda ainsi défini au début de ce siècle. Encore plus nombreux sont ceux parmi nous qui ont connu cette période vers le milieu des années 2000 pendant laquelle Gerdi Gerschheimer animait à Paris un séminaire où indianistes et sud-est asiatisants, philologues et archéologues, chercheurs chevronnés et débutants, se rencontraient autour de la lecture d'inscriptions sanskrites du Cambodge.

Nous devons sans doute à la dynamique créée par le CIK le fait que notre École a pu concourir en 2018 avec succès pour un projet européen d'envergure, de type « Synergy », qui crée un partenariat avec le Centre d'Études de l'Inde et d'Asie du Sud, un laboratoire CNRS-EHESS, ainsi qu'avec les universités Humboldt et L'Orientale, respectivement à Berlin et à Naples. Ce projet DHARMA, acronyme pour « The development of Hindu 'asceticism' and the religions making of South and Southeast Asia », indique par son titre que nous nous intéressons à l'histoire sociale et religieuse sur l'ensemble des mondes indien et indianisé. Mais ce titre ne dit pas que le projet est pluridisciplinaire, car il comprend des fouilles archéologiques au Bangladesh, en Indonésie et au Cambodge ; il ne dit pas non plus que plusieurs éditions de textes, sanskrits ou vieux-javanais, transmis sur feuilles de palmier, seront réalisées dans son cadre ; et le titre ne dit pas, enfin, que le cœur du projet est la création d'une vaste base de données épigraphiques rassemblant des corpus linguistiquement hétérogènes, provenant de divers pays (Inde, Népal, Bangladesh, Cambodge, Thaïlande, Vietnam, Indonésie) et couvrant une période d'environ quinze siècles à partir du début de notre ère.

Comment en sommes-nous arrivés, à partir du « Corpus des inscriptions khmères », à pouvoir envisager une base aussi vaste ? Il faut évoquer premièrement le fait que Daniel Perret avait lancé en 2002 un « Programme franco-indonésien d'inventaire des inscriptions "classiques" du monde malais »[7]. Ce programme a vu en 2021 paraître son site web sous le titre IDENK, ou « Inventaire en ligne de l'épigraphie de l'ancien nousantara »[8]. Deuxièmement, dans le sillage du programme lancé par Gerdi Gerschheimer sur le pays khmer, j'avais moi-même lancé en 2009 un programme de « Corpus des inscriptions du Campā », ou CIC, pour relancer l'engagement de notre École dans ce domaine épigraphique. C'est dans le cadre du CIC que nous avons commencé à expérimenter en 2011 un travail d'édition

7. D. Perret *et al.*, 2003-2004.

8. *Inventaris Daring Epigrafi Nusantara Kuno / Online Inventory of Ancient Nusantaran Epigraphy*. idenk.net.

selon les normes de la « Text Encoding Initiative », ou TEI. Il s'agit là d'un système de « recommandations pour la création et la gestion sous forme numérique de tout type de données créées et utilisées par les chercheurs en sciences humaines », et cela « indépendamment de tout environnement logiciel particulier »[9]. Plus précisément, nous avons commencé à travailler selon la mise en application des normes TEI élaborées par des spécialistes de l'épigraphie classique sous le nom d'EpiDoc[10]. Grâce à l'aide précieuse d'un des concepteurs d'EpiDoc, Tom Elliott, a vu le jour en 2012 ce qui fut probablement le premier site web consacré à un corpus d'épigraphie asiatique encodé selon les normes TEI-EpiDoc[11]. Je reviendrai dans quelques instants sur l'intérêt de ce cadre informatique, mais termine d'abord le survol des programmes qui ont ouvert le chemin vers le projet DHARMA, en mentionnant les sites web « Early Inscriptions of Āndhradeśa » (EIAD)[12] et « Corpus of Pyu Inscriptions » (PYU)[13], que nous avons pu créer entre 2015 et 2017. Il s'agit là, du côté indien, d'un corpus d'inscriptions relevant de l'histoire du bouddhisme dans le sud de l'Inde pendant les premiers siècles de notre ère, le plus souvent rédigées en moyen-indien. De l'autre côté du golfe du Bengale, en Birmanie, il s'agit d'un corpus associé aux grands sites classés sur la Liste du patrimoine mondial de l'UNESCO depuis 2014. Les corpus EIAD et PYU ont été réalisés en collaboration avec Emmanuelle Morlock du laboratoire HiSoMA à Lyon, ingénieure experte en TEI-EpiDoc.

Les avantages de ce système de balisage résident avant tout dans son caractère sémantique, qui fait totalement abstraction de considérations de mise en forme d'une édition, et permet d'outrepasser les limites de recherches sur des simples chaînes de caractères dans des éditions saisies numériquement en Word ou texte brut. Chaque élément d'une édition, qu'il s'agisse d'une lettre, d'un mot, d'une phrase, ou de quelque autre élément que ce soit, peut être balisé pour ce qu'il signifie[14]. Le corpus DHARMA englobe des

9. L. Burnard, 2015, p. 1.

10. https://epidoc.stoa.org/gl/latest/intro-intro-fr.html.

11. *Corpus of the Inscriptions of Campā*. https://isaw.nyu.edu/publications/inscriptions/campa/.

12. http://hisoma.huma-num.fr/exist/apps/EIAD/index2.html.

13. http://hisoma.huma-num.fr/exist/apps/pyu/index2.html.

14. Voir les remarques de H. Cayless *et al*., 2009, § 26, à propos du système dit « Leiden » de conventions éditoriales commune dans l'épigraphique classique, et de l'utilité de les représenter en TEI-EpiDoc : « By the term "queryable", we do not simply mean that the text may be scanned for particular patterns of characters; we mean that features of the text indicated by Leiden should be able to be investigated also. So, for example, a corpus of inscriptions should be able to be queried for the full list of abbreviations used within it, or for the number of occurrences of a word in its full form, neither abbreviated nor supplemented. »

cultures épigraphiques qui n'ont jamais encore été considérées ensemble, quoiqu'il y ait consensus sur leur parenté. Dès lors que les inscriptions seront insérées dans ce système éditorial constituant un cadre analytique cohérent, nous serons en mesure de questionner ces sources primaires sous des angles qui dépassent les cloisonnements géographiques, chronologiques, linguistiques, nationaux ou autres, qui empêchent actuellement de saisir les caractéristiques communes à ces diverses cultures épigraphiques et celles qui les singularisent[15].

Permettez-moi de donner trois exemples, en commençant par la passionnante chasse aux premières occurrences du chiffre « zéro ». Nous savons tous que la notion de « zéro » a été conçue en Inde, mais quand et où a-t-on commencé à le mettre par écrit sous la forme d'un petit rond ? L'idée que l'honneur d'avoir conservé le plus ancien exemple d'un chiffre zéro écrit reviendrait au Cambodge, et que cet exemple daterait de la fin du VII^e^ siècle, circule depuis quelques années[16], sans mention systématique du fait que le chiffre zéro figure à la même époque dans des inscriptions de l'île de Sumatra et sans que compte soit tenu des observations pertinentes de Louis-Charles Damais, dans les pages de notre *Bulletin*[17]. Grâce au fait que toutes nos éditions sont balisées à la fois pour leurs dates de gravure en ère commune, pour leurs provenances géographiques, et pour tous les chiffres décimaux (ou autres) qu'ils contiennent, il nous sera possible de faire une carte qui montre quand et où le chiffre zéro est noté dans les inscriptions. Voici un bout de code, avec la visualisation correspondante, pour une inscription du royaume de Śrīvijaya, datée de 684 de notre ère, qui montre comment on prépare les données pour permettre la réponse à ce genre de questionnement (fig. 5). Ce sont les balises, ici en bleu, ajoutées autour des phrases et de certains mots, qui indiquent à l'ordinateur qu'il s'agit d'annotations insérées par l'éditeur du texte. Ainsi par exemple que

15. Voir la présentation par Irene Vagionakis (2021, § 9) des avantages du modèle TEI-EpiDoc : « EpiDoc's semantic nature allows users to mark up efficiently not only the traditional components of a critical edition of an inscription (such as its metadata, text, apparatus, translation, commentary, and bibliography), but also specific interpretive features occurring in the epigraphic texts themselves, which could be linguistic, historical, topographic, prosopographical, and so on. The flexibility of the EpiDoc encoding, moreover, allows the digital epigraphists to choose what features to mark up, according to their particular research topics. A collection of semantically encoded inscriptions, in fact, could be queried in a very accurate way, allowing digital epigraphists to identify recurring patterns and features more easily. Furthermore, EpiDoc markup could lead to faster indexing of specific elements, such as places or personal names, that are mentioned in the texts. »

16. Par exemple, Diller, 1995 et Diller, 1996.

17. Voir Damais, 1952, p. 98-99, n. 1 et 2.

```
<lb n="1"/>||  • śrī śakavarṣātīta • <num value="606">606</num> diṁ
dvitīya śuklapakṣa vulan· caitra • sāna tatkālāña parlak· śrīkṣetra Ini • niparvuAt·
<lb n="2"/>parvāṇḍa punta hiyaṁ śrī <persName type="human"
subtype="coronation">jayanāga</persName> • Ini praṇidhānāṇḍa punta hiyaṁ ...
```

⑴|| • śrī śakavarṣātīta • 606 diṁ dvitīya śuklapakṣa vulan· caitra • sāna tatkālāña parlak· śrīkṣetra Ini • niparvuAt·
⑵parvāṇḍa punta hiyaṁ śrī jayanāga • Ini praṇidhānāṇḍa punta hiyaṁ ...

FIG. 5. – Exemple (a) de code TEI-EpiDoc pour les deux premières lignes d'une inscription illustrant les balises <lb>, <num> et <persName>, avec (b) la « mise en page » du même passage.

```
<lg n="1" met="jagaddhita" xml:lang="kaw-Latn"><!-- The second pāda has deviant fourth gaṇa. -->
        <l n="a">saṁ hyaṁ viṣṇu sirāśarīra sira riṁ bhuvana subhagavastha riṁ prajā,</l>
        <l n="b" real="+++--+-+--++------+-+-+">svastha śrī jayavarṣa-digjayaśāstraprabhu saphala sinəmbah iṁ sarāt·,</l>
        <l n="c">sākṣāt· bhaskaracandra tīrtha sira tāmṛta ri hajeṅ ikaṁ sarāt· kabeḥ<supplied reason="lost">,</supplied></l>
        <l n="d">Astvāniṅky i sahaśracandra paṅadəgnira sinivi haneṁ jagat·kṛtā //</l>
</lg>
```

Jagaddhita

saṁ hyaṁ viṣṇu sirāśarīra sira riṁ bhuvana subhagavastha riṁ prajā,
svastha śrī jayavarṣa-digjayaśāstraprabhu saphala sinəmbah iṁ sarāt·,
sākṣāt· bhaskaracandra tīrtha sira tāmṛta ri hajeṅ ikaṁ sarāt· kabeḥ[,]
Astvāniṅky i sahaśracandra paṅadəgnira sinivi haneṁ jagat·kṛtā//

FIG. 6. – Exemple (a) de la représentation en TEI-EpiDoc d'une strophe en vieux javanais illustrant les balises <lg>, <l> et <supplied>, ainsi que la fonction commentaire, avec (b) la « mise en page » de la strophe.

le texte est disposé sur deux lignes interrompu par un retour à la ligne, <lb n="1"/> et <lb n="2"/>, ou qu'une suite de caractères représente un nombre <num> dont la valeur numérique, explicitée dans l'attribut « value », en orange, est 606. Cette démarche est essentielle pour que la machine puisse repérer et trier les caractères sur lesquels elle doit effectuer une opération.

Mon deuxième exemple tient au fait que la structure métrique de versification sanskrite a été adoptée au-delà de l'Inde et de son application à la langue sanskrite, un peu à la manière dont l'hexamètre dactylique a été utilisé tant pour la poésie métrique en grec que pour la versification latine. Ainsi, c'est notamment sur l'île de Java que la vaste gamme de structures métriques d'origine indienne a été redéployée pour donner forme à une littérature en langue locale. Si maintenant nous voulions avoir des données sur la diffusion dans le temps et dans l'espace de certaines formes métriques, tout en gardant le contrôle sur le contexte linguistique – trouver par exemple seulement les occurrences javanaises d'un type de mètre donné –, ce sera la combinaison des balises pour la typologie métrique et celles pour la langue dans laquelle une stance est rédigée qui permettra d'obtenir des résultats significatifs (fig. 6). Ici, les attributs en orange indiquent que la strophe <lg>

portant le numéro 1 est en mètre Jagaddhita, en langue kaw, ce qui signifie le vieux javanais[18], et qu'elle est représentée en caractères latins, indiqués par le code -Latn[19]. On indique également, tant pour le lecteur humain (en vert) que pour l'ordinateur, avec l'attribut « real », que le second vers de la strophe est en fait construit sur un schéma métrique de syllabes courtes et longues différent de celui qui est attendu dans le cas d'une strophe *Jagaddhita*, et enfin, par le biais de la balise <supplied>, qu'un signe de ponctuation a dû être restauré par l'éditeur.

Mon troisième et dernier exemple concerne le phénomène de diffusion lexicale entre le pays Tamoul et le Sud-Est asiatique, et la question de l'abréviation de termes techniques. Comparons ces deux passages rédigés respectivement en tamoul et en vieux javanais (fig. 7) :

```
<measure unit="kaḻañcu" quantity="100" commodity="gold">nūṟṟu-k kaḻañcu poṉṉum</measure>
<measure unit="kāṭi" quantity="28" commodity="paddy">Irupatteṇ kāṭi nellum</measure> kuṭuttēṉ

InaṁsəAn· pasak-pasak· <measure quantity="1" unit="kāṭi" commodity="gold" >mas·
<choice><abbr>kā</abbr><expan>kāṭi</expan></choice> <num value="1">1</num></measure>
<measure quantity="1" unit="yugala" commodity="vḍihan taṅkalan">vḍihan· taṅkalan·
<choice><abbr>yu</abbr><expan>yugala</expan></choice> <num value="1">1</num></measure>
```

nūṟṟu-k kaḻañcu poṉṉum Irupatteṇ kāṭi nellum kuṭuttēṉ

InaṁsəAn· pasak-pasak· mas· kā 1 vḍihan· taṅkalan· yu 1

Fig. 7. – Exemple (a) de la représentation en TEI-EpiDoc d'un passage d'une inscription tamoule ainsi que d'une inscription javanaise illustrant les balises <measure>, <choice>, <abbr>, <expan> et <num>, avec (b) la « mise en page » des deux passages.

Le premier signifie : « J'ai donné une centaine de *kaḻañcu* en or et vingt-huit *kāṭi* en paddy », le second : « On lui a offert des dons à raison de 1 *kāṭi* en or et 1 *yugala* en tissu (de type) *taṅkalan* ». Le balisage permet non seulement de faire apparaître que l'emploi de l'unité de poids *kāṭi* est commun aux corpus tamoul et javanais, mais aussi quelles marchandises étaient quantifiées en *kāṭi*. Il est en outre intéressant de constater dans ces deux exemples que le javanais favorise les abréviations <abbr> d'unités et les quantités en chiffres <num>, tandis que le tamoul favorise les termes complets et les quantités en toutes lettres. Ceci nous amène à nous poser la question de la généralisation d'un tel constat.

On voit donc que le balisage qualifié de « sémantique » enrichit les données et accroît leur potentiel d'exploitation sous de multiples angles

18. https://iso639-3.sil.org/code/kaw.
19. https://unicode.org/iso15924/iso15924-codes.html.

scientifiques, ce que je n'ai pu montrer que trop brièvement ici. C'est une évidence que le fait d'enrichir les données de cette façon nécessite un travail qui va bien au-delà du déchiffrement et de la traduction des textes épigraphiques. Pareil investissement de temps peut devenir un frein. Néanmoins, l'avantage d'œuvrer dans un cadre numérique est de pouvoir procéder par couches. On peut se contenter dans un premier temps d'un balisage relativement simple, et appliquer par la suite des couches successives d'annotation. Autrement dit, on peut se servir de TEI-EpiDoc pour publier un corpus sans avoir procédé à tous les types de balisage que j'ai illustrés.

Un autre défi pour parvenir à exploiter le riche potentiel du balisage en TEI-EpiDoc est le besoin d'ingénierie compétente pour accompagner l'ensemble de la chaîne informatique. Nous nous sommes heurtés au fait de ne pas pouvoir bénéficier de toute l'ingénierie nécessaire lors de la création des corpus CIC, EIAD et PYU, ce qui explique pourquoi ces sites web sont restés dans un état expérimental. C'est ensuite, dans le cadre du projet DHARMA, qu'il est devenu possible d'embaucher trois ingénieur.e.s, un nombre qui n'est guère luxueux au vu de la quantité et de la complexité du travail à accomplir. Un défi pour l'avenir sera d'assurer la pérennité des données tandis que l'EFEO ne dispose pas en son sein d'un ingénieur qui puisse se consacrer à la tâche[20].

J'ai insisté jusqu'à présent sur ce que le tournant vers le numérique implique pour le travail épigraphique une fois que les documents originaux ont été déchiffrés. Mais les apports du numérique pour la documentation des artefacts inscrits, sur le terrain ou dans des musées, ne sont pas moindres. Je me limite à illustrer ici les nouvelles méthodes de reproduction qui viennent en complément de l'estampage, sans par ailleurs que ce dernier soit devenu moins pertinent, ainsi que les perspectives passionnantes qu'offre la numérisation des fonds patrimoniaux d'estampages, conservés dans des collections publiques en France ou ailleurs.

Pour commencer donc avec les nouvelles méthodes de reproduction, nous travaillons de plus en plus souvent avec les technologies de la photogrammétrie et de la « Reflectance Transformation Imaging », soit RTI. Vous voyez ici l'exemple d'une inscription rupestre du Nord du Cambodge. Son déchiffrement s'annonce pénible. En revanche, la reproduction en

20. Voir I. Vagionakis, 2021, § 21 : « My previous experience, however, also showed me that it is not always possible to transform these potentialities into reality quickly, especially in the context of small projects lacking the support of IT experts. And wider epigraphic projects often benefit from IT support only in the initial and final phases of the work. In some cases, the epigraphists are able to see the fruits of their intense and years-long encoding effort only at the publication stage, at the end of their work, which may be quite discouraging! »

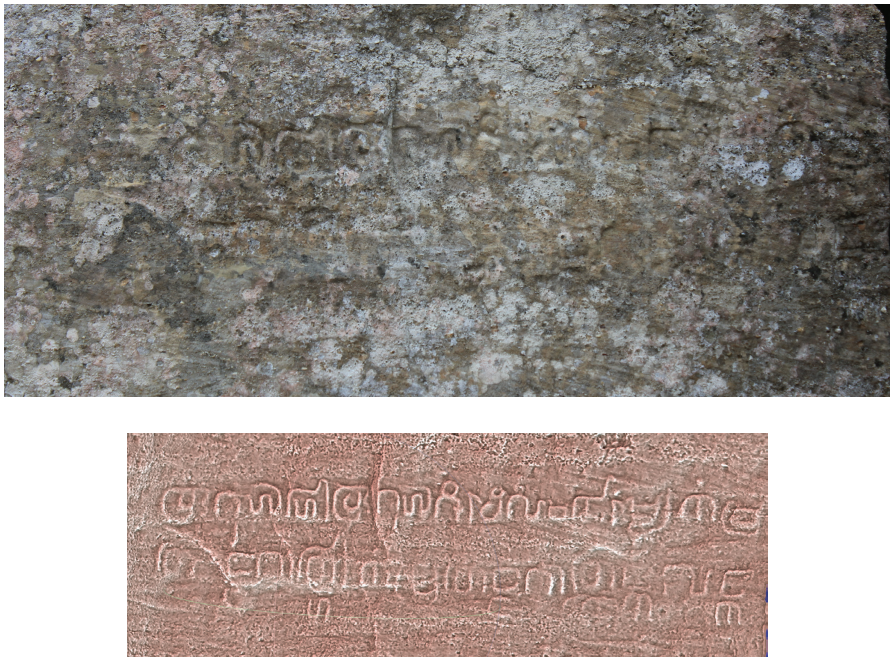
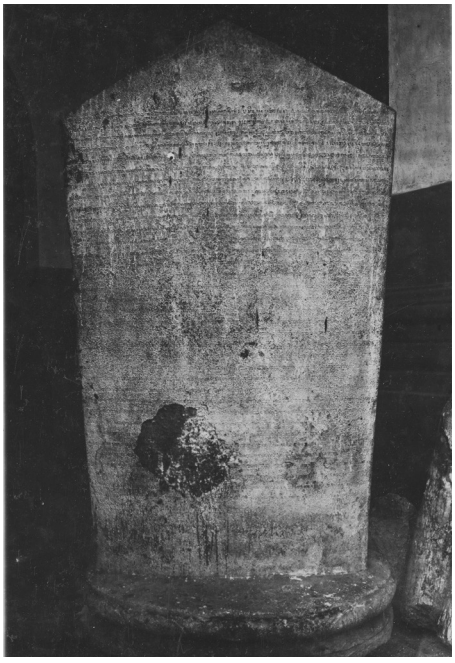


Fig. 8. – Inscription K. 1495, Phnom Tbeng, province de Preah Vihear, Cambodge, (a) photographie et (b) orthophotographie à partir d'un modèle 3D (cliché Chloé Chollet 2021 ; photogrammétrie Nicolas Josso 2021).

photogrammétrie, restitue à cette inscription sa lisibilité (fig. 8). Les résultats peuvent être tout aussi réjouissants dans le cas de recours à la RTI. Comparez cette photographie de piètre qualité en haut avec le joli rendu violet qu'offre la visualisation RTI en bas (fig. 9). La disponibilité de ces nouvelles formes d'imagerie ne permet pas de réparer tous les dommages qu'ont subis les inscriptions au fil des siècles, mais peut souvent rendre un peu moins ardu le travail de déchiffrement. De ce point de vue, le fait que nous ayons aujourd'hui un accès infiniment plus facile au terrain, pour aller lire les inscriptions *in situ* ou dans les musées d'Asie directement sur leurs supports physiques, nous donne un avantage fort considérable par rapport à nos prédécesseurs qui devaient se contenter d'accéder aux inscriptions à travers les estampages.

Mais les cas ne sont pas rares où des inscriptions connues autrefois ont été perdues, ou ont souffert de dégâts importants depuis leur découverte au XIX^e^ ou au XX^e^ siècle. C'est donc une chose particulièrement heureuse que la plupart des estampages réalisés à l'époque coloniale aient été conservés du côté français dans la bibliothèque de notre École, dans celle de la Société Asiatique, et à la Bibliothèque nationale. Du côté néerlandais, un bon



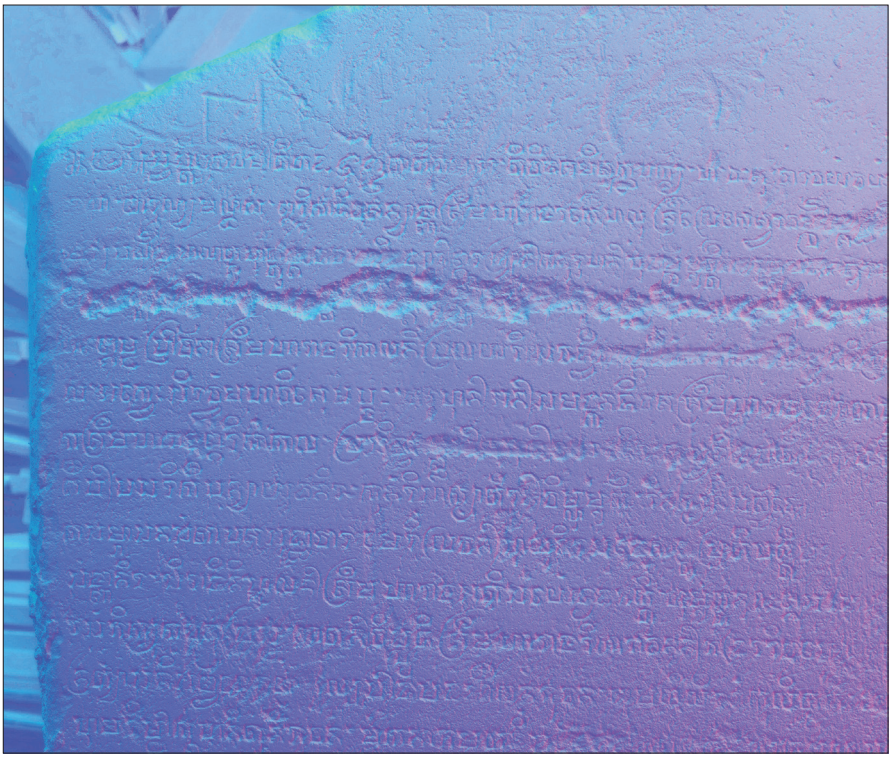


Fig. 9. – Inscription de Pucangan, Java Est (a) photographie et (b) extrait de RTI (cliché EFEO INDO00163, anonyme, 1951 ; RTI réalisé par Archaeovision, 2019).

Fig. 10. – Estampages (a) de la BnF (fonds Inscriptions du Campā et du Cambodge, carton 46, nº 422) et (b) de la bibliothèque universitaire de Leyde, fonds Institut Kern, D55 (clichés Adeline Levivier, 2019-2021).

nombre sont conservés dans la bibliothèque de Leyde. Les prises de vue avec lumière rasante permettent de donner des rendus forts satisfaisants même dans le cas d'estampages non encrés (fig. 10). Le fait que de plus en plus d'établissements soient convaincus de devoir laisser entièrement libre l'accès aux images de documents qu'ils conservent nous ouvre la voie à l'inclusion dans la base de données DHARMA des images que nous avons réalisées à Paris comme à Leyde, et nous pouvons espérer que les très importantes collections de l'Archaeological Survey of India, à Mysore (fig. 11), puissent également devenir accessibles un jour sous forme numérique[21].

Après 120 ans d'existence de l'EFEO, nous qui y poursuivons l'œuvre des fondateurs de l'épigraphie de l'Asie du Sud-Est, visons aujourd'hui à profiter de façon optimale de notre accès privilégié au terrain et des nouvelles méthodes de reproduction pour réaliser l'édition des inscriptions avec la plus exigeante des rigueurs philologiques. Notre activité est également marquée par un engagement profond dans les humanités numériques, dans le but de permettre à ces sources précieuses de nous dire tout ce qu'elles peuvent sur le passé de l'Asie.

Arlo Griffiths

Le projet DHARMA est financé de 2019 en 2025 par le Conseil Européen de la Recherche (ERC) dans le cadre du programme de recherche et d'innovation de l'Union Européenne Horizon 2020 (contrat de financement nº 809994).

21. Voir aussi la conférence de Michèle Brunet (2021) présentant la plate-forme E-stampages et illustrant le fait que le contexte de publication et l'outillage numériques (TEI-EpiDoc ou la plate-forme documentaire sur les estampages) ne sont pas spécifiques à telle ou telle langue ou culture, mais sont des « génériques » créés comme des systèmes éditoriaux et des moyens communs pour toute une discipline.

Fig. 11. – Documentation d'estampages pour le corpus EIAD
à l'Archaeological Survey of India, Mysore (cliché Adeline Levivier, 2017).

*
* *

Références bibliographiques

Michèle Brunet, 2021, « E-stampages : la mise en ligne des collections d'estampages. Une nouvelle ressource pour l'étude des inscriptions grecques », *Comptes rendus des Séances de l'Académie des Inscriptions et Belles-Lettres* 2019, fasc. I (janv.-mars), p. 263-285.

Lou Burnard, 2015, *Qu'est-ce que la Text Encoding Initiative ?* URL : http://books.openedition.org/oep/1237. Consulté le 28 septembre 2022.

Hugh Cayless, Charlotte Roueché, Tom Elliott *et al.*, 2009, « Epigraphy in 2017 », *Digital Humanities Quarterly* 3/1.

George Cœdès, 1908, « Inventaire des inscriptions du Champa et du Cambodge », *Bulletin de l'École française d'Extrême-Orient* 8, p. 37-92.

—, 1918, « Le royaume de Çrīvijaya », *Bulletin de l'École française d'Extrême-Orient* 18/6, p. 1-36.

—, 1923, « Liste générale des inscriptions du Champa et du Cambodge », in *Listes générales des inscriptions et des monuments du Champa et du Cambodge*, Hanoi, EFEO, p. 11-80.

George Cœdès, 1930, « Les inscriptions malaises de Çrīvijaya », *Bulletin de l'École française d'Extrême-Orient* 30/12, p. 29-80.

—, 1937, *Inscriptions du Cambodge I*, Hanoi, EFEO (Collection de textes et documents sur l'Indochine, 3.1).

—, 1966, *Inscriptions du Cambodge VIII*, Paris, EFEO (Collection de textes et documents sur l'Indochine, 3.8).

Louis Charles Damais, 1952, « Études d'épigraphie indonésienne, III : Liste des principales inscriptions datées de l'Indonesie », *Bulletin de l'École française d'Extrême-Orient* 46/1, p. 1-105.

Anthony Diller, 1995, « Sriwijaya and the First Zeros », *Journal of the Malaysian Branch of the Royal Asiatic Society* 68/1 (268), p. 53-66.

—, 1996, « New Zeros and Old Khmer », *Mon-Khmer Studies* 25, p. 125-132.

Louis Finot, 1915, « Notes d'épigraphie, XX : L'épigraphie indochinoise », *Bulletin de l'École française d'Extrême-Orient* 15/2, p. 113-135.

—, « Notes d'épigraphie, I : deux nouvelles inscriptions de Bhadravarman I^{er}, roi de Champa », *Bulletin de l'École française d'Extrême-Orient* 2, p. 185-191.

Louis Finot et George Coedès, 1926-1937, *Inscriptions du Cambodge publiées sous les auspices de l'Académie des Inscriptions et Belles-Lettres*, Paris, Geuthner.

Gerdi Gerschheimer, 2003-2004, « Le Corpus des inscriptions khmères », *Bulletin de l'École française d'Extrême-Orient* 9091, p. 478-482.

Bahour S. Kuppusamy, 2006, *Pondicherry inscriptions. Part I: Introduction and texts with notes*, éd. G. Vijayavenugopal, Pondichéry, EFEO-IFP.

—, 2010, *Pondicherry inscriptions. Part II: Translation, appendices, glossary and phrases, [and] preface*, éd. G. Vijayavenugopal, Pondichéry, EFEO-IFP.

Daniel Perret, Machi Suhadi et Richadiana Kartakusuma, 2003-2004, « Le programme franco-indonésien d'inventaire des inscriptions "classiques" du monde malais », *Bulletin de l'École française d'Extrême-Orient* 90-91, p. 473-474.

Charlotte Schmid, 2005, « Au seuil du monde divin : Reflets et passages du dieu d'Ālanturai à Puḷḷamaṅkai », *Bulletin de l'École française d'Extrême-Orient* 92, p. 39-157.

Pullur Ramasubrahmanya Srinivasan et Marie-Louise Reiniche, 1990, *Tiruvannamalai: A Śaiva sacred complex of South India. Vol. 1: Inscriptions*, Pondichéry, IFP (Publications de l'Institut français d'Indologie, 2 vol., n° 75).

Irene Vagionakis, 2021, « Cretan Institutional Inscriptions: A New EpiDoc Database », *Journal of the Text Encoding Initiative* (septembre 2021). URL : http://journals.openedition.org/jtei/3570. Consulté le 29 novembre 2021.