Consignes liminaires pour la sélection des images.

Critères retenus pour la sélection :

* Critères architecturaux / caractères externes de l’inscription.

Hypothèse A de Matthieu : on peut faire une typologie monumentale à partir d’un type de textes. Ex : textes de dédicace, peuvent être sur monument A, monument B, mais pas monument C.

Donc partir de la table des matières « list of text Categories », tenter de dresser une typologie architecturale pour chaque typologie textuelle.

Hypothèse B de Matthieu : la machine serait capable d’identifier la provenance du lieu de l’édifice au sein du complexe d’Aphrodisias (dans les métadonnées : *findspot*), en fonction : inscription du stade, du théâtre, etc.

Hypothèse B de Thomas : A partir du texte, reconnaissance du texte de l’inscription 🡺 reconnaissance du support architectural de l’inscription ; un pool entraînement et un pool libre pour voir s’il fonctionne.

On part donc de la table des matières « List of monuments ».

Vu qu’une même typo d’inscription peut se retrouver sur différents supports monumentaux, on aura un système de probabilités (ex : frise 70 %, architrave 25 %, autel 5%) ; et si hésitation prolongée, faire porter son attention sur les caractères externes pour combler son hésitation.

Corpus hypothèse A Matthieu :

Attention : items may have multiple categories.

Acclamation : retenue

Agonistic : retenue

Alphabet : non retenu (1 item)

Building dedication / building inscription : retenues

Column dedication : retenu (15 item)

Consolatory decree : non retenu (7 item)

Consolary text : non retenu (1 item)

Curse : non retenu (9 item)

Decree : retenu

Dedication + dedication to demos + dedication to emperor doctor : retenu

Donor’s inscription : non retenu (9 item)

Edict : non retenu (1 item)

Factional (sens ? 🡪 lutte de factions ? Associations d’artistes) : retenu

Foundation : non retenu (<15 item)

Funerary + funerary verse : retenu

Gameboard : non retenu (<15)

Gladiator + gladiator memorial : retenu (mais nbx recoupements avec acclamations)

Graffiti + graffito : retenu (mais typologie textuelle branlante)

Honours : retenu.

Imperial + imperial letter : retenu.

Insults : non retenu (<15)

Invocation : non retenu (<15)

Jewish : non retenu (<15)

Letter : retenu

List : non retenu (<15)

Mason’s mark : non retenu (<15)

Measurement : non retenu (<15)

Monogram : non retenu (<15)

Munera : non retenu (<15)

Names : non retenu (<15)

Numbering : non retenu (<15)

Oath : non retenu (<15)

Pavement design : non retenu (<15)

Place inscription : retenu.

Posthumous honors : retenu (recoupements avec honours)

Prayer + prayer / invocation : retenu

Prohibition : non retenu (<15)

Regulation : non retenu (<15)

Religious : retenu (mais catégorie textuelle étrange)

Report : non retenu (<15)

Restoration : non retenu (<15)

Sculptor’s signature : non retenu (<15)

Seat inscriptions : retenu

Senatus consultum : non retenu (<15)

Statue dedication + statue dedication of aphrodite : retenu (recoupements avec dedication)

Subscript : non retenu (<15)

Sundial : non retenu (<15)

Testament : non retenu (<15)

Title : retenu

Verse : retenu

Votive : non retenu (<15)

Faire la différence entre les inscriptions classées par sujet (religieux) et les inscriptions classées par type (ex. acclamation).

Eviter principe de non contradiction, accueillir et utiliser. Faire des statistiques sur corrélation. Intuitivement, des réponses.

Corpus avec 18 classes, au moins 1500 images ; mais scrap = 4 000 images.

Si on pose une question historique, mail à personnes tenant la base de données, obtient la baseline pour savoir à quoi on est confronté. Permet de savoir ce qui a été fait, etc. Si ne répondent pas, tant pis.