

Humanités Numériques

Cours 3 : Linguistique et Informatique

Loïc Grobol

2021-09-21

L'index thomiscus

<https://www.corpusthomisticum.org/it/index.age>

De quoi s'agit-il ?

Le travail de Roberto Busa

Roberto Busa (1913 — 2011) était un prêtre jésuite italien spécialiste de Thomas d'Aquin

According to the scholarly practices, I first > searched through tables and subject indexes for the words of *praesens* and *praesentia*

present

. I soon learned that such words in Thomas Aquinas are peripheral : his doctrine of presence is linked with the preposition *in*. My next step was to write out by hand 10 000 3“×5” cards, each containing a sentence with the word *in* or a word connected with *in*. Grand games of solitaire follow. ((Busa, 1980))

-
- L'objectif initial de Busa est philosophique et théologique
 - Mais sa thèse est qu'on ne peut accéder à la pensée d'un auteur que si on maîtrise sa façon d'employer le langage.
 - Il lui faut donc étudier Thomas d'un point de vue philologique et linguistique.
 - Il formule donc en 1946 le projet d'un grand concordancier des œuvres de Thomas

-
- Réalisation immédiate : la tâche est trop vaste pour être faite sans assistance
 - Il se met donc en quête de « machinerie » pour l'aider

any gadget that might help ((Busa, 1980))

- Il parvient à obtenir l'aide gracieuse d'IBM pour le réaliser
- Le projet a duré 30 ans, pour produire une transcription complète de 179 en forme lisible par des machines l'époque (des cartes perforées !), et une lemmatisation semi-automatique.

-
- 10 631 980 mots

- 1 500 km de câble
- 10 000 h de calcul
- 1 000 000 d'heures de travail humain

On identifie souvent l'*index thomiscus* comme le premier travail d'humanités numériques mené à terme, même si Busa suggère que d'autres initiatives plus anciennes ont existé.

Bibliographie

R. Busa. 1980. The Annals of Humanities Computing : The Index Thomisticus. *Computers and the Humanities*, 14(2):83-90.