

16|17 NOVEMBRE
2016

CITÉ DES SCIENCES ET
DE L'INDUSTRIE - PARIS



Les rencontres du
NUMÉRIQUE
de l'ANR

ANR
10
ANS

La Traduction Automatique

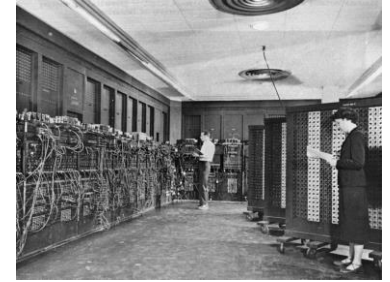
Quelques repères

Deux manières de traduire

- **Qualité**
- **Délais**
- **Coût**
 - Expertise rare
 - Formation longue



- **Instantanéité**
- **Disponibilité**
- **Versatilité**
- **Coût**
- **Qualité**



Une hybridation nécessaire

Traduction Outillée

- Mémoires de traduction
- Pré/post-édition
- Traduction interactive/prédictive
- Traduction en réseau

- Les interfaces du traducteur

The screenshot shows the Matecat web interface. At the top, there's a header with the Matecat logo and a document title: "Referentiels_compétences...lccompagnementLicence.pdf (353202) > fr-FR > en-GB". There are buttons for "PREVIEW" and a chat icon. The main content area is titled "La démarche en questions" and "The process in question". It shows a French text on the left and its English translation on the right. A blue box highlights the English translation: "Why repeat the process of developing a competency framework for undergraduate studies?". Below the text, there are tabs for "Translation Matches", "Concordance", and "Glossary". The "Translation Matches" tab is active, showing a table of matches:

Source	Target	Score
Pourquoi reprendre la démarche d'élaboration d'un référentiel de compétences pour les licences ?	Why repeat the process for developing a competency framework for licensing?	Source: MT
Référentiel de compétences	Skills repository	Source: Public TM 2015-05-19 47%
référentiel de compétences	competency framework	Source: Public TM 2013-02-18 47%

Below the table, there are several paragraphs of text in French, with the first one partially visible: "Le nouveau contexte réglementaire introduit par le cadre national des formations et l'importance que doivent revêtir le supplément au diplôme ainsi que l'inscription des diplômés au répertoire national des certifications professionnelles (RNCP), engage à réinterroger et compléter les travaux déjà réalisés sur les référentiels de compétences en licence."

Brève histoire de la TA

TA 1960-1990



- 1- Systèmes experts + règles linguistiques
- 2- *Logiciellisation* de la TA
- 3- La TA est « *AI – complete* »

TA 1990-



- 1- Apprentissage des ressources
- 2- Apprentissage des composants
- 3- Apprentissage global

Des données pour apprendre

- Traductions humaines
- Textes en langue cible

En grande quantité

De grande qualité

GERMAN

Einleitung

I. Von dem Unterschiede der reinen und empirischen Erkenntnis

Daß alle unsere Erkenntnis mit der Erfahrung anfangt, daran ist gar kein Zweifel; denn wodurch sollte das Erkenntnisvermögen sonst zur Ausübung erweckt werden, geschähe es nicht durch Gegenstände, die unsere Sinne rühren und teils von selbst Vorstellungen bewirken, teils unsere Verstandstätigkeit in Bewegung bringen, diese zu vergleichen, sie zu verknüpfen oder zu trennen, und so den rohen Stoff sinnlicher Eindrücke zu einer Erkenntnis der Gegenstände zu verarbeiten, die Erfahrung heißt? Der Zeit nach geht also keine Erkenntnis in uns vor der Erfahrung vorher, und mit dieser fängt alle an.

ENGLISH

Introduction

I. Of the difference between Pure and Empirical Knowledge

That all our knowledge begins with experience there can be no doubt. For how is it possible that the faculty of cognition should be awakened into exercise otherwise than by means of objects which affect our senses, and partly of themselves produce representations, partly rouse our powers of understanding into activity, to compare to connect, or to separate these, and so to convert the raw material of our sensuous impressions into a knowledge of objects, which is called experience? In respect of time, therefore, no knowledge of ours is antecedent to experience, but begins with it.

FRENCH

Introduction

I. De la différence de la connaissance pure et de la connaissance empirique.

Que toute notre connaissance commence avec l'expérience, cela ne soulève aucun doute. En effet, par quoi notre pouvoir de connaître pourrait-il être éveillé et mis en action, si ce n'est par des objets qui frappent nos sens et qui, d'une part, produisent par eux-mêmes des représentations et, d'autre part, mettent en mouvement notre faculté intellectuelle, afin qu'elle compare, lie ou sépare ces représentations, et travaille ainsi la matière brute des impressions sensibles pour en tirer une connaissance des objets, celle qu'on nomme l'expérience? Ainsi, chronologiquement, aucune connaissance ne précède en nous l'expérience et c'est avec elle que toutes commencent.

Apprendre à traduire

Modéliser : $t \hat{=} T, s \hat{=} S : t = f_q(s)$

Questions: Choix de $f()$, des représentations de s et t (+/- analysés, des arbres syntaxiques / sémantiques)

Estimer: $\{(t_i, s_i), i = 1 \dots N\} \rightarrow q^* = \arg \min L(D; q)$

Questions: Choix de L / doit assurer $f_q(s_i) \gg t_i$; métrique (comment évaluer =) ?

Inférer: $t^* = f_q(s)$

Questions: Coût du calcul, incrémentalité etc.

Etat de l'art

Modèles neuronaux $f_q(s) = \arg \max \text{NN}_q(s, t)$

performances, adaptation, temps d'apprentissage, opacité

Représentations « simplistes » : mots, groupes de lettres, caractères

multilinguisme, vocabulaires ouverts

Calcul incrémental de s|t

vitesse de traitement, latence