



Productions participatives de corpus annotés : des modèles encore incertains

Karën Fort

► To cite this version:

Karën Fort. Productions participatives de corpus annotés : des modèles encore incertains. Colloque Jeunes Chercheurs PRAXILING, Nov 2019, Montpellier, France. hal-02418826

HAL Id: hal-02418826

<https://inria.hal.science/hal-02418826>

Submitted on 19 Dec 2019

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.



Productions participatives de corpus annotés : des modèles encore incertains

Karën Fort

karen.fort@sorbonne-universite.fr

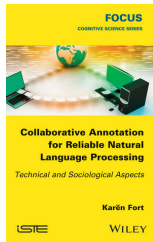
8 novembre 2019



D'où je parle

Voir <http://karenfort.org/>

- Création de ressources langagières pour le Traitement automatique des langues (TAL)



- Ethique et TAL



Production participative (*crowdsourcing*)

Jeux ayant un but que j'ai participé à créer :



ZOMBILINGO

RIGORMORTIS

BISAME

KRIK !

AYO !



Portail de jeux pour les langues et atelier récurrent :



Games4NLP

Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

Conclusion et perspectives

Les productions participatives

Des participants qu'on ne choisit pas

Au-delà des mythes

Redéfinition

Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

Conclusion et perspectives

Production participative / myriadisation (*crowdsourcing*)

Crowdsourcing is *"the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call."* [Howe, 2006]

- ▶ pas d'identification ou de sélection des participants *a priori* (appel ouvert)
- ▶ massif (en termes de production et de participation)
- ▶ (relativement) peu cher

Des réussites remarquables

Wikipédia¹ (octobre 2019) :

- ▶ plus de **158 million d'articles** en **279** langues
- ▶ plus de 800 millions de pages de la Wikipédia française vues en septembre

Distributed Proofreaders (Gutenberg Project)² :

- ▶ près de **40 000** livres numérisés et corrigés

Numérisation des déclarations de conflits d'intérêts des élus³ :

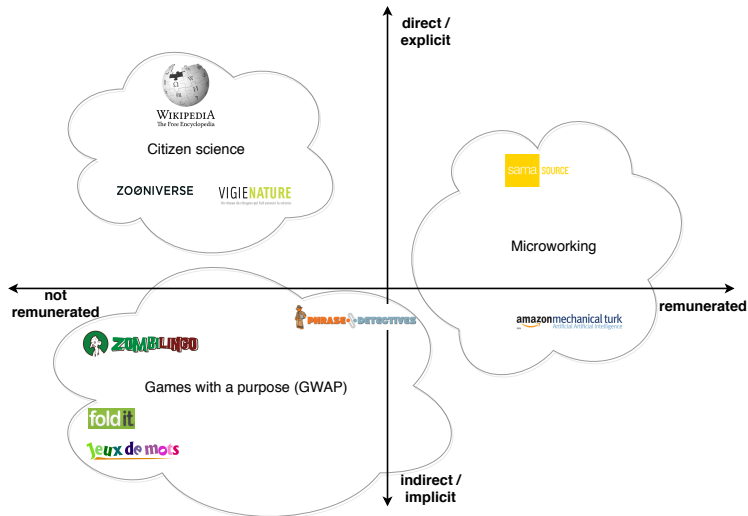
- ▶ 11 095 extraits de déclarations saisis **en moins d'une semaine**
- ▶ près de 8 000 participants

1. <https://stats.wikimedia.org/v2/#/all-wikipedia-projects>

2. <https://www.pgdp.net/c/>

3. <http://regardscitoyens.org/interets-des-elus/>

Les productions participatives



Les productions participatives

Des participants qu'on ne choisit pas

Au-delà des mythes

Redéfinition

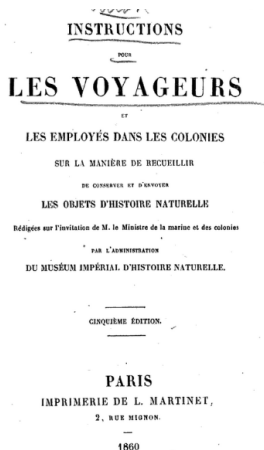
Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

Conclusion et perspectives

Mythe #1 : « C'est nouveau ! »

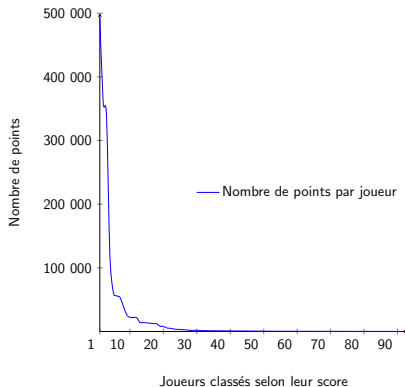
Instructions pour les voyageurs et les employés des colonies



Sciences participatives :

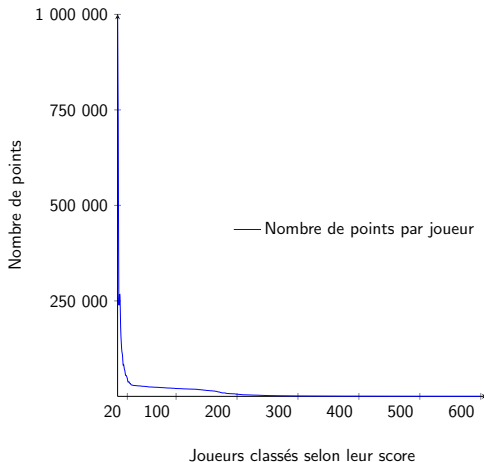
- ▶ publiées par le Museum National d'Histoire Naturelle
- ▶ première édition en **1824**

Mythe #2 : « Ca implique beaucoup de participants »



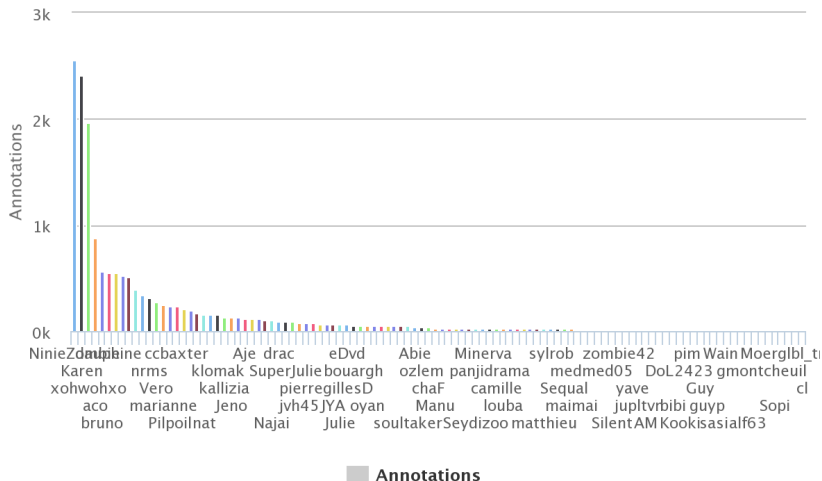
Nombre de joueurs sur *Phrase Detectives* selon le nombre de points gagnés dans le jeu (fév. 2011 - fév. 2012)

Une foule de participants ? JeuxDeMots



Nombre de joueurs sur *JeuxDeMots* selon leur classement dans le jeu
(source : <http://www.jeuxdemots.org/generateRanking-4.php>)

Une foule de participants ? ZombiLingo



Highcharts.com

Une foule de travailleurs ? [Fort et al., 2011]

Nombre de *Turkers* actifs sur Amazon Mechanical Turk :

- ▶ annoncé : plus de 500 000
- ▶ 80 % des tâches (HIT) sont réalisés par les 20 % des *Turkers* les plus actifs [Deneme, 2009]
- ⇒ travailleurs vraiment actifs (en 2011) : entre 15 059 et 42 912

Mythe #3 : « Ca implique des non-experts »

Extraits du forum de ZombiLingo

4 zombies ont donné leur avis

1084 (Phrase de référence)



25 %

Le principal critère d'évaluation de l'efficacité a été de mesurer si le taux dans le sang de phosphatases alcalines sériques (enzyme impliquée dans la dégradation des os) est revenu à la normale ou **est** **REDESCENDU** d'au moins 75 % pour se rapprocher des taux normaux .

Suivre la discussion

Discuter de la réponse

Annotation expert

Justin a écrit il y a 9 mois :

est redescendu est le passé composé du verbe redescendre, de même que "est revenu" est le passé-composé du verbe revenir. Si on les considérait comme des présents de l'indicatif au passif, il faudrait pouvoir écrire "il a été revenu" et "il a été redescendu". Ce qui n'est pas le cas....

17 zombies ont donné leur avis

22637



35 %

Graphes associés à un jeu de Nim Dans la grande majorité des cas , l'ensemble des situations , même s'il est très grand , **est** **FINI** .

Suivre la discussion

Discuter de la réponse

Annotation expert

Zeltron a écrit il y a 9 mois :

l'ensemble est fini donc il n'est pas infini : rien à voir avec finir

Les productions participatives

Des participants qu'on ne choisit pas

Au-delà des mythes

Redéfinition

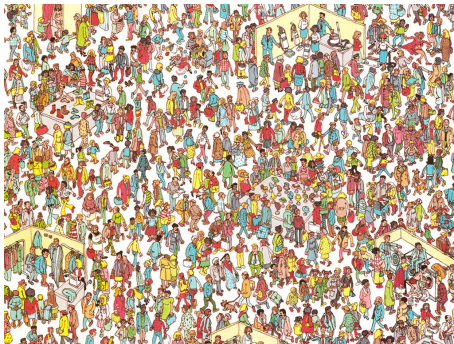
Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

Conclusion et perspectives

Production participative

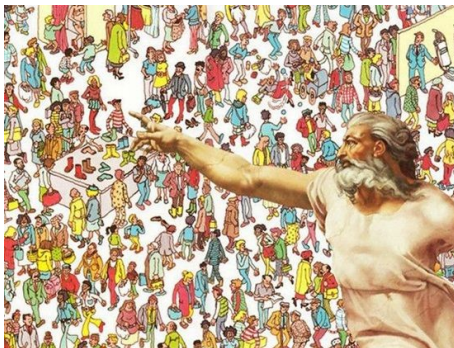
Profiter d'une foule de "non-experts" ?



Production participative

~~Profiter d'une foule de "non-experts" ?~~

→ Trouver/former des experts (de la tâche) dans la foule



Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

- Les jeux ayant un but

- La syntaxe de dépendances, en jeu

- Derrière le rideau

- Résultats

Les incertitudes du modèle

Conclusion et perspectives

JeuxDeMots : jouer à faire des associations d'idées. . .

. . . pour créer un réseau lexical [Lafourcade and Joubert, 2008]

Près de 300 million de relations (créées par env. 6 000 joueurs),
mises à jour constamment

- ▶ jeu par pairs
- ▶ des relations de plus en plus complexes, typées
- ▶ des challenges
- ▶ des procès
- ▶ etc



Phrase Detectives : jouer les détectives...

... pour annoter des anaphores [Chamberlain et al., 2008]

3,5 millions de décisions de
45 000 joueurs (2016)

- ▶ corpus pré-annoté
- ▶ instructions détaillées
- ▶ formation
- ▶ 2 modes de jeu :
 - ▶ annotation
 - ▶ validation (correction d'annotations)

DETECTIVES CONFERENCE

Another detective has made a decision about a phrase, either that it refers to another phrase, it has not been mentioned before, it is a property or it does not refer to anything. Do you agree with them?

Knitta (Wikipedia)

Polycotl and Molyk came up with their own names, then invented names for other members in a brainstorming session they considered "one of the more hilarious meetings". Some former member names include Knotious N.I.T., SonOfStich and P-Kinky.

As of January 2008, the group has two female members and one male, ages 30 to 73, who wish to remain anonymous. Current members' are Polycotl, MaccuKitty, and Granny SQ. An estimated five to twelve frequent groups exist around the world.

Usually tagging on Friday nights and Sunday mornings, **Knitta** members leave a paper tag on each work, leaving the slogan "Knitta please" or "Whaddup knitta?". They tag trees, lamp posts, railings, fire hydrants, monuments and other urban targets, and even get a little "hardcore" with ideas like hanging knitted bagged smokers over aerial telephone cables. The crew marks holidays by doing themed work, using, for example, pink yarn for their Valentine's Day pieces and sparkly yarn for New Years. When Knitta is not working with a theme, they work on projects, tagging specific targets or specific areas.

The group and their followers consider their graffiti "a method of beautifying public space".

The phrase in blue is the **closest** phrase that refers to the phrase in orange.

Disagree Agree

NAME THE CULPRIT

Has the phrase shown in orange been mentioned before in this list or is it a property? Use your mouse to select the **closest phrase(s)** if it has been mentioned before.

Knitta (Wikipedia)

Polycotl and Molyk came up with their own names, then invented names for other members in a brainstorming session they considered "one of the more hilarious meetings". Some former member names include Knotious N.I.T., SonOfStich and P-Kinky.

As of January 2008, the group has two female members and one male, ages 30 to 73, who wish to remain anonymous. Current members' are Polycotl, MaccuKitty, and Granny SQ. An estimated five to twelve frequent groups exist around the world.

Usually tagging on Friday nights and Sunday mornings, Knitta members leave a paper tag on each work, bearing the slogan "Knitta please" or "Whaddup knitta?". They tag trees, lamp posts, railings, fire hydrants, monuments and other urban targets, and even get a little "hardcore" with ideas like hanging knitted bagged smokers over aerial telephone cables. The crew marks holidays by doing themed work, using, for example, pink yarn for their Valentine's Day pieces and sparkly yarn for New Years. When Knitta is not working with a theme, they work on projects, tagging specific targets or specific areas.

Knitta. Play work on projects, tagging specific targets or specific areas.

Not mentioned before This is a property

Done

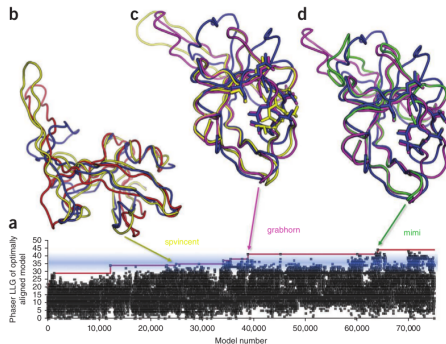
FoldIt : jouer à replier des protéines. . .

. . . pour résoudre des problèmes de structure cristalline [Khatib et al., 2011]

Résolution de la structure cristalline de la protéine responsable de la propagation du virus du SIDA chez les macaques rhésus

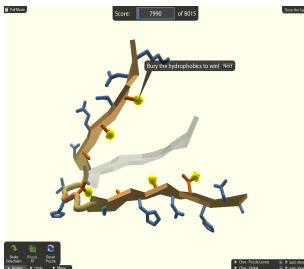
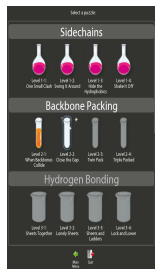
Solution à un problème non résolu depuis plus de 10 ans

- ▶ trouvée en quelques semaines
- ▶ par une équipe de joueurs
- ▶ qui va permettre la création de médicaments antirétroviraux



FoldIt : jouer à replier des protéines. . .

... sans aucune connaissance préalable en biochimie [Cooper et al., 2010]



Formation par étapes

- ▶ tutoriel décomposé par concepts
- ▶ puzzles pour chaque concept
- ▶ l'accès aux puzzles suivants n'est octroyé que si le niveau du joueur est suffisant

Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les jeux ayant un but

La syntaxe de dépendances, en jeu

Derrière le rideau

Résultats

Les incertitudes du modèle

Conclusion et perspectives

Une tâche complexe



- ▶ guide d'annotation
 - ▶ 29 types de relation
 - ▶ approx. 50 pages
- ▶ des décisions contre-intuitives (**pas** de la grammaire d'écoliers, de la linguistique) : aobj = au

[...] avoir recours au type de mesures [...]

c-à-d que la tête de la relation est ici une préposition

→ décomposer la complexité de la tâche [Fort et al., 2012],
pas la simplifier !

ZOMBI LINGO

BIENVENUE à TOI, JEUNE ZOMBIE !

LE MONDE EST
CONDMANÉ, TA
TRANSFORMATION EN
ZOMBIE A COMMENCE.

POUR SURVIVRE,
SUIS MES RÈGLES,
IDENTIFIE LES TÊTES
ET MANGE-LES.

ATTENTION AUX
PIÈGES, ILS SONT
NOMBREUX !



SIGNE

Prof. Frankensperre.

Jouer

*Pas de limite
pour toi !
Tu accèdes à
toutes les
options, bonus
cachés !*

Karen

*Retrouve ici tes
statistiques, et
compare ton
score avec
celui de tes
amis !*

SEMAINE

MOIS
TOTAL

1 nouveau : 38 993
2 nicolef : 16 170

3 bruno : 11 546
4 testKF : 7 327
5 Karen : 7 012



Jouer

Boutique

Forum

Joueurs



Niveau

maximum!

165

Trouve le complément (objet indirect introduit par "à") du verbe indiqué !

10%

Besoin
d'aide?



Très jeune, il a fait preuve d'initiative et de courage pour
PARTICIPER à un sauvetage lors d'inondations.



57

150

Acheter

1

15

Acheter



3

300

Acheter

74



0

15

Acheter

43

Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les jeux ayant un but

La syntaxe de dépendances, en jeu

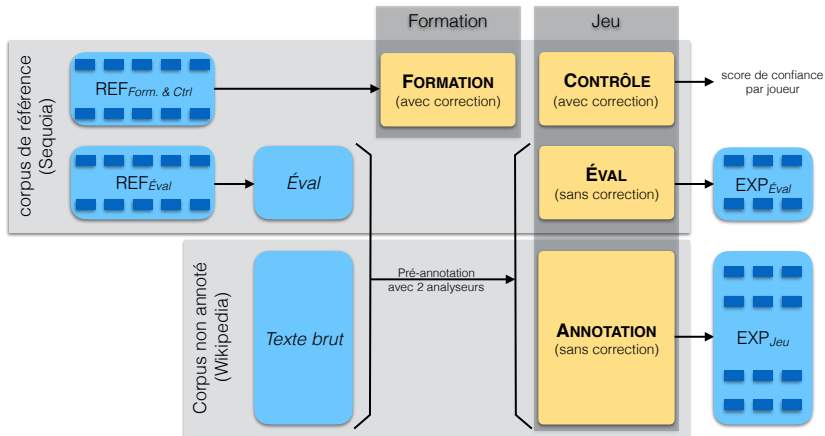
Derrière le rideau

Résultats

Les incertitudes du modèle

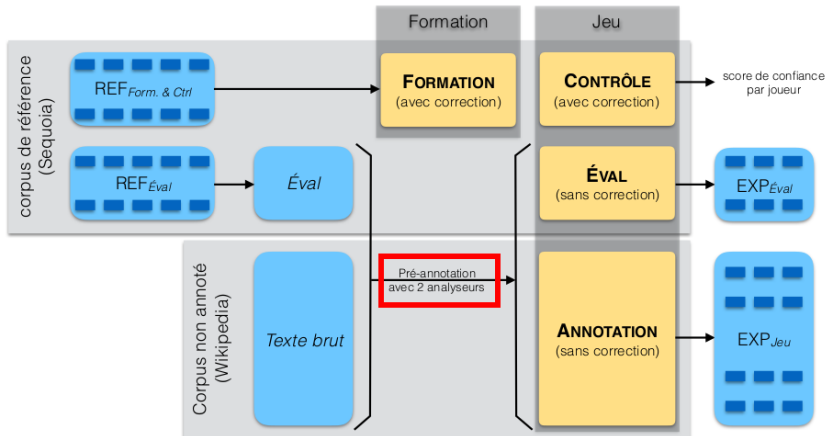
Conclusion et perspectives

Organiser une production de qualité



Prétraitement des données

corpus librement disponibles et distribuables



Prétraitement des données

corpus librement disponibles et distribuables

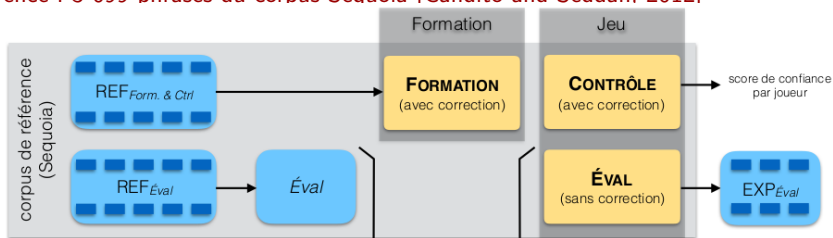
Pré-annotation avec deux parsers

1. un statistique : Talismane [Urieli, 2013]
2. un symbolique, basé sur la ré-écriture de graphes :
FRDEP-PARSE [Guillaume and Perrier, 2015]

→ les joueurs ne jouent que les relations sur lesquelles les deux parsers ne donnent pas les mêmes résultats

Formation, contrôle et évaluation

Référence : 3 099 phrases du corpus Sequoia [Candito and Seddah, 2012]



REF _{Train&Control}	REF _{Eval}	non utilisé
50 %	25 %	25 %
1 549 phrases	776 phrases	774 phrases

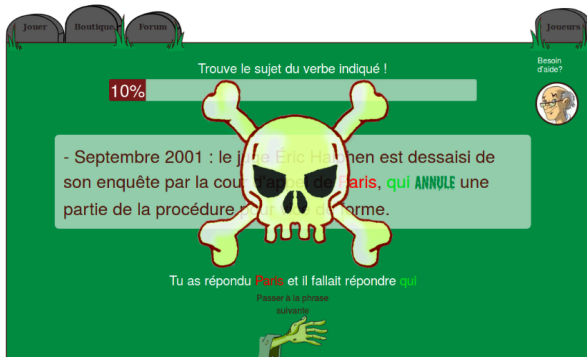
- ▶ REF_{Train&Control} est utilisé pour former les joueurs
- ▶ REF_{Eval} est utilisé comme un corpus brut, pour évaluer les annotations produites

Formation



Obligatoire pour chaque relation

- ▶ phrases du corpus REF_{Train&Control}
- ▶ retour visuel en cas d'erreur



Gestion de la fatigue cognitive et des joueurs au long court

Mécanisme de contrôle

Des phrases de REF_{Train&Control} sont proposées régulièrement

1. si le joueur échoue à trouver la bonne réponse, un retour visuel avec la solution lui est proposé

Ils ont été reçus à la boulangerie Leroy **POUR visiter** le fournil et **surtout** pétrir la pâte afin de confectionner de délicieux pains au chocolat qu'ils ont dégustés à l'heure du goûter avec un verre de jus de fruit.

Tu as répondu **surtout** et il fallait répondre **visiter**

Il te reste 2 essais avant de devoir refaire le tutoriel de ce phénomène

⚠ Je ne suis pas d'accord

Passer à la phrase
suivante



Gestion de la fatigue cognitive et des joueurs au long court

Mécanisme de contrôle


Des phrases de REF_{Train&Control} sont proposées régulièrement

1. si le joueur échoue à trouver la bonne réponse, un retour visuel avec la solution lui est proposé
2. après un certain nombre d'erreurs sur une même relation, le joueur ne peut plus jouer et doit refaire la formation correspondante

- 1er **FÉVRIER** 1995 : Jean-Paul Schimpf, un ami intime de Didier Schuller, est **arrêté** sur un parking, alors que la dirigeante d'une entreprise d'assainissement disait vouloir lui remettre **une** somme d'argent en liquide.

Tu as répondu **une** et il fallait répondre **arrêté**

Tu as un peu oublié comment jouer ce phénomène. Pour continuer à jouer sur celui-ci, tu vas devoir refaire le tutoriel correspondant.

 Je ne suis pas d'accord

[Retourner au menu](#)



Gestion de la fatigue cognitive et des joueurs au long court

Mécanisme de contrôle

Des phrases de REF_{Train&Control} sont proposées régulièrement

1. si le joueur échoue à trouver la bonne réponse, un retour visuel avec la solution lui est proposé
 2. après un certain nombre d'erreurs sur une même relation, le joueur ne peut plus jouer et doit refaire la formation correspondante
- nous en déduisons un **niveau de confiance** dans le joueur, pour **cette** relation

Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les jeux ayant un but

La syntaxe de dépendances, en jeu

Derrière le rideau

Résultats

Les incertitudes du modèle

Conclusion et perspectives

Production : taille des corpus créés

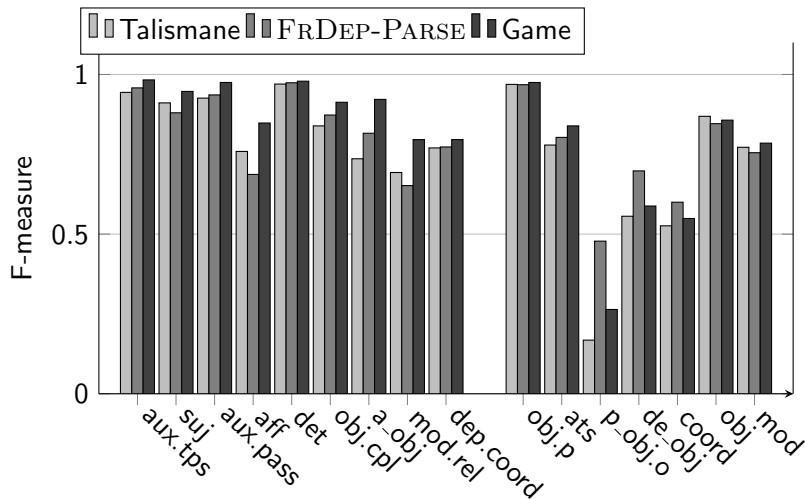
Au 10 juillet 2016

- ▶ 647 joueurs (1 460 au 9 octobre 2019)
- ▶ ont produit 107 719 annotations (496 462 au 9 octobre 2019)

→ ressource qui évolue constamment !

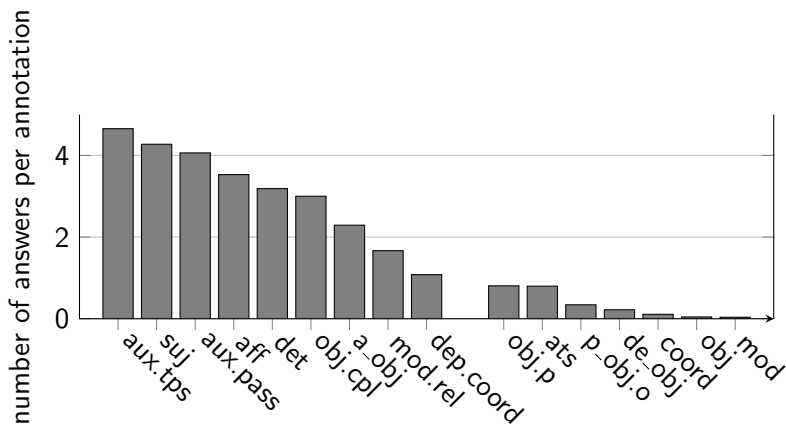
Évaluation de la qualité

sur le corpus REF_{Eval}



Densité des annotations

sur le corpus REF_{Eval}



→ besoin de **davantage** d'annotations sur certaines relations

Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

- Créer un jeu : un savoir-faire complexe

- Produire des données de qualité

- Motiver la participation : de l'extrinsèque vers l'intrinsèque

Conclusion et perspectives

Types de joueurs selon [Bartle, 1996]

Motiver les joueurs en fonction de leur type

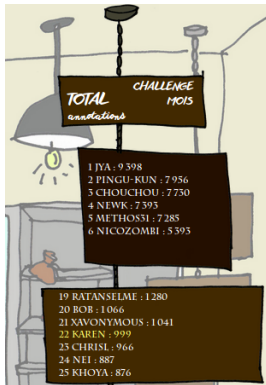
- ▶ *Achievers* : aiment réussir dans le jeu
- ▶ *Explorers* : aiment connaître les coins cachés du jeu
- ▶ *Socializers* : aiment interagir avec les autres
- ▶ *Killers* : aiment attaquer les autres joueurs

Classements (pour tout le monde)



A cartoon illustration of a scoreboard hanging from a ceiling. The scoreboard has two sections: 'TOTAL points' and 'CHALLENGE MOIS'. The 'TOTAL points' section lists the top 16 players and their scores. The 'CHALLENGE MOIS' section lists the top 6 players and their scores. The background shows a desk with a lamp and a small figure sitting at it.

TOTAL points	CHALLENGE MOIS
1 JYA : 5 406 318	
2 PINGU-KUN : 3 417 699	
3 CHOUCOU : 2 387 725	
4 XOHWOHXO : 2 378 725	
5 DIHTHI : 2 222 847	
6 ZOMBINOCLARD : 2 180 338	
10 LYCO : 745 379	
11 BRUNO : 710 177	
12 ... : 604 591	
13 KAREN : 509 753	
14 NINIEZOMBIE : 564 932	
15 IZNONGOOD : 283 346	
16 BOB : 265 772	



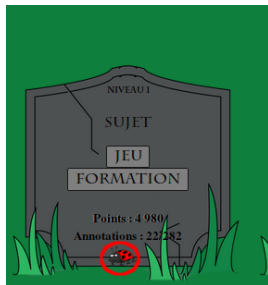
A cartoon illustration of a scoreboard hanging from a ceiling. The scoreboard has two sections: 'TOTAL annotations' and 'CHALLENGE MOIS'. The 'TOTAL annotations' section lists the top 25 players and their scores. The 'CHALLENGE MOIS' section lists the top 6 players and their scores. The background shows a desk with a lamp and a small figure sitting at it.

TOTAL annotations	CHALLENGE MOIS
1 JYA : 9 398	
2 PINGU-KUN : 7 956	
3 CHOUCOU : 7 730	
4 NEWK : 7 393	
5 METHOSSY : 7 285	
6 NICOZOMBIE : 5 393	
19 RATANSELMIE : 1 280	
20 BOB : 1 066	
21 XAVONYMOUS : 1 041	
22 KAREN : 999	
23 CHRISL : 966	
24 NEI : 887	
25 KHOYA : 876	

Critères :

- ▶ nombre d'annotations ou de points
- ▶ au total, au moins, pendant le défi

Fonctionnalités cachées (pour les explorateurs)



- ▶ qui apparaissent aléatoirement
- ▶ différents effets : apparition d'objet, lien vers un autre jeu, etc.

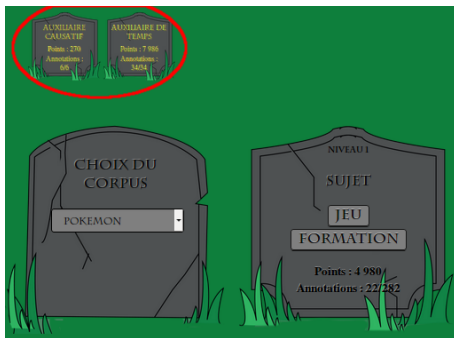
Duels (pour les "socialiseurs" (et les tueurs?))

The screenshot shows a game interface with a green background. At the top, there are five buttons: JOUER, BOUTIQUE, DUELS, FORUM, and ADMIN. On the right side, there is a button labeled JOUEURS. Below the buttons, the word "Duels" is displayed next to a "Nouveau duel" button. On the right, there is a summary of duel statistics: "Duels gagnés : 1", "Duels perdus : 1", and "Matches nuls : 1". Below this, there are three tabs: "DUELS LIBRES 1", "MES DUELS EN COURS 2", and "MES DUELS TERMINÉS 3". The "MES DUELS EN COURS 2" tab is selected. Below the tabs, there is a table of duels. The table has four columns: a status icon, the names of the players and their scores, the result of the duel, and a description of the duel.

	DUELS LIBRES 1	MES DUELS EN COURS 2	MES DUELS TERMINÉS 3	
🔥	Karen vs. JYA 1420 1558	JYA gagne 5956 points	Complément en « de » duel en 20 en tours	
	Karen vs. bruno 660 660	Karen gagne 660 points bruno gagne 660 points	Auxiliaire causatif duel en 20 en tours	
👍	Karen vs. JYA 1754 1670	Karen gagne 3424 points	Complément en « de » duel en 20 en tours	

- ▶ sélectionner un ennemi
- ▶ lui proposer un duel sur un type de relation

Badges (?) (pour les collectionneurs)



- ▶ jouer toutes les phrases pour une relation
- ▶ jouer toutes les phrases pour un corpus

Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

Créer un jeu : un savoir-faire complexe

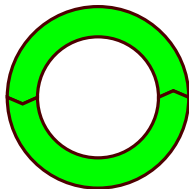
Produire des données de qualité

Motiver la participation : de l'extrinsèque vers l'intrinsèque

Conclusion et perspectives

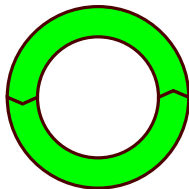
Données de qualité vs Fonctionnalités fun

préserver le cercle vertueux n'est pas si facile



Données de qualité vs Fonctionnalités fun

préserver le cercle vertueux n'est pas si facile

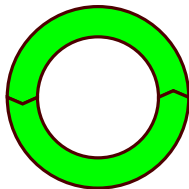


phrase qui disparaît soudainement dans ZombiLingo :

- + le joueur est surpris : fun !
- le joueur clique n'importe où : ressource de mauvaise qualité

Données de qualité vs Fonctionnalités fun

préserver le cercle vertueux n'est pas si facile



phrase qui disparaît soudainement dans ZombiLingo :

- + le joueur est surpris : fun !
- le joueur clique n'importe où : ressource de mauvaise qualité

joueur qui a trouvé une faille dans JeuxDeMots [Lafourcade and Joubert, 2008] pour obtenir du temps :

- + crée de la meilleure donnée : bonne qualité
- génère de l'envie et de la colère dans la communauté de joueurs : mauvais pour le jeu

Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

Créer un jeu : un savoir-faire complexe

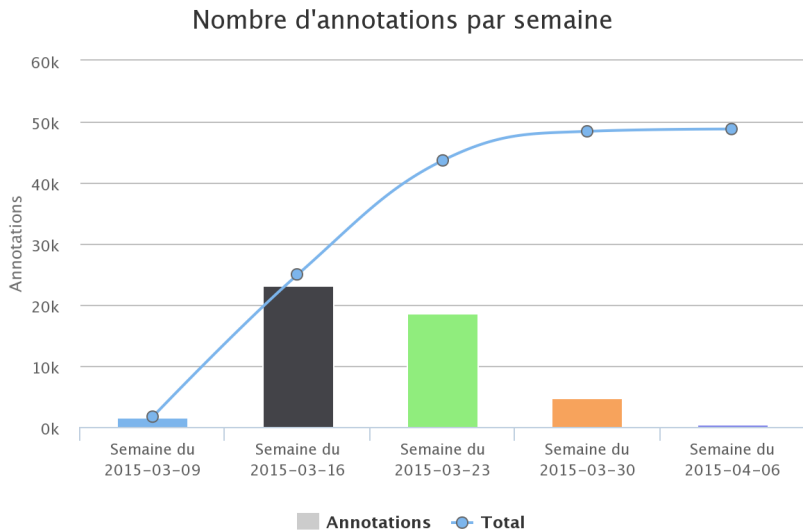
Produire des données de qualité

Motiver la participation : de l'extrinsèque vers l'intrinsèque

Conclusion et perspectives

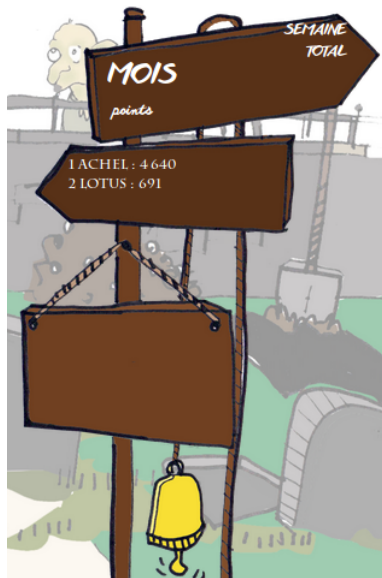
Extrinsèque : à relancer constamment

Exemple de ZombiLingo (influence de la semaine de la langue française)



Highcharts.com

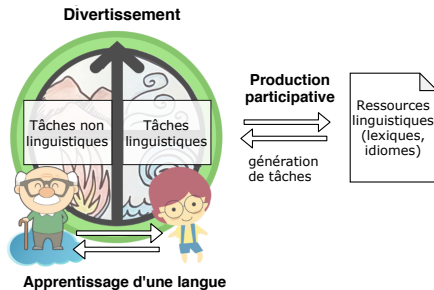
Quand on ne relance pas



Intrinsèque : apprentissage d'une langue

favoriser la transmission inter-générationnelle en produisant des données

Katana et Grand Guru, le jeu des mots perdus



Les productions participatives

Les jeux ayant un but : l'exemple de ZombiLingo

Les incertitudes du modèle

Conclusion et perspectives

Jouir de la foule est un art [Baudelaire, « Les foules »]

Faire réaliser des tâches complexes nécessite de :

- ▶ connaître les dimensions de complexité de la tâche
- pour outiller à bon escient
- ▶ former les joueurs et les évaluer
- pour les sélectionner

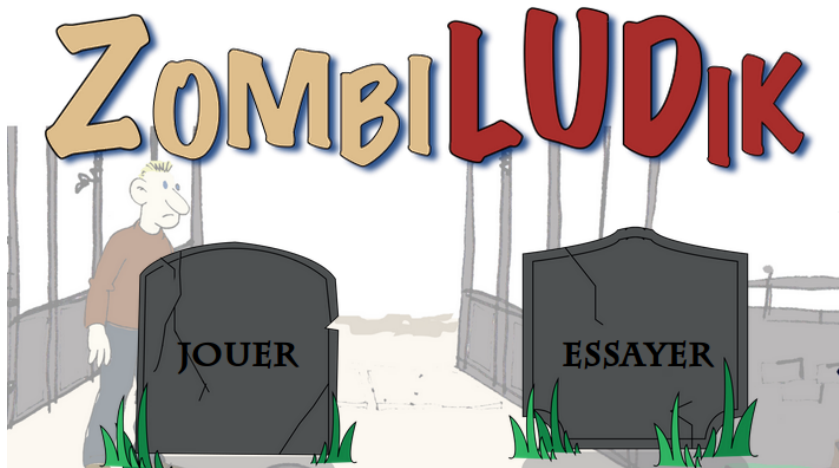
Obtenir des ressources de qualité nécessite de :

- ▶ déterminer les moyens et les formes de l'évaluation
- ▶ créer un cercle vertueux



Bientôt sur vos écran !

Le jeu pour Universal Dependencies : <http://zombiludik.org/>



THANK  YOU



<https://github.com/zombilingo>



<http://zombilingo.org/export>

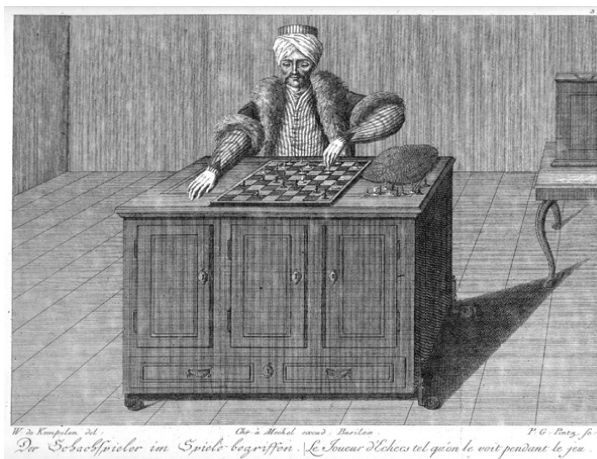
Appendix

Amazon Mechanical Turk : une plate-forme de légendes

Bibliographie

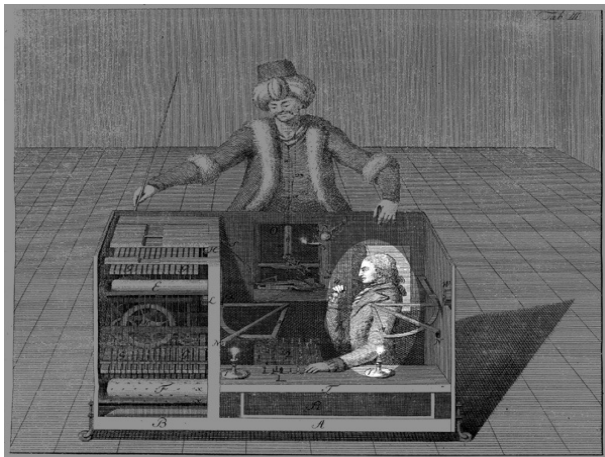
Le «Turc mécanique» de von Kempelen

Un joueur d'échecs mécanique créé par J. W. von Kempelen en 1770 :



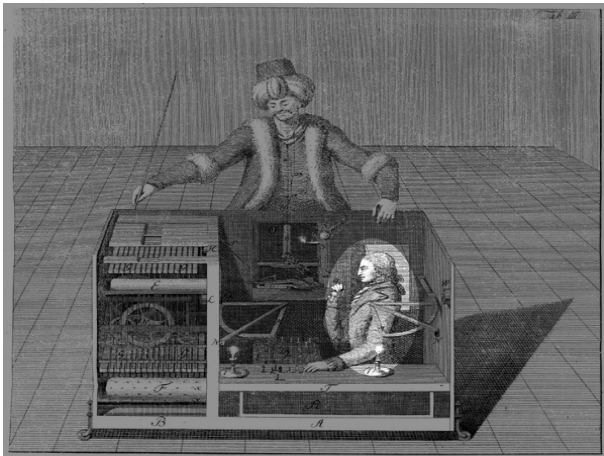
Le «Turc mécanique» de von Kempelen

En fait, un maître d'échecs était caché dans la machine :



Le «Turc mécanique» de von Kempelen

C'est l'intelligence artificielle **artificielle** !



Et Amazon créa AMT

Amazon crée une pour ses propres besoins
plate-forme de travail parcellisé
et en ouvre l'accès en 2005 (moyennant 10 % des transactions)

Amazon Mechanical Turk

MTurk

[Your Account](#)[HITs](#)[Qualifications](#)

Already have an account?
Sign in as a [Worker](#) | [Requester](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Mechanical Turk is a marketplace for work.

We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.

179,373 HITs available. [View them now.](#)

Make Money by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work



Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Register Now](#)


As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results



Amazon Mechanical Turk

MTurk est une plate-forme de **myriadisation** : le travail est *externalisé* via le Web et réalisé par de nombreuses personnes (la *foule*), ici les **Turkers**



[Your Account](#)[HITS](#)[Qualifications](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Mechanical Turk is a marketplace for work.

We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient.

179,373 HITS available. [View them now.](#)

Make Money by working on HITS

HITS - Human Intelligence Tasks - are individual tasks that you work on. [Find HITS now.](#)


As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task

Work

Earn money



Get Results from Mechanical Turk Workers

Ask workers to complete HITS - Human Intelligence Tasks - and get results using Mechanical Turk. [Register Now](#)


As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITS completed in minutes
- Pay only when you're satisfied with the results

Fund your account


Load your tasks

Get results



Amazon Mechanical Turk

MTurk est une plate-forme de **myriadisation** du travail parcellisé : les tâches sont découpées en sous-tâches (HIT) et leur exécution est payée par les **Requesters**



[Your Account](#)[HITS](#)[Qualifications](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Amazon Mechanical Turk
Artificial Intelligence

Already have an account?
[Sign in as a Worker](#) | [Requester](#)

Mechanical Turk is a marketplace for work.

We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient.

179,373 HITS available. [View them now.](#)


Make Money by working on HITS

HITS - Human Intelligence Tasks - are individual tasks that you work on. [Find HITS now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task → **Work** → **Earn money**



[Find HITS Now](#)


Get Results from Mechanical Turk Workers

Ask workers to complete HITS - Human Intelligence Tasks - and get results using Mechanical Turk. [Register Now](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITS completed in minutes
- Pay only when you're satisfied with the results


Fund your account → **Load your tasks** → **Get results**



[Get Started](#)

Amazon Mechanical Turk

MTurk est une plate-forme de **myriadisation** du travail parcellisé : les tâches sont découpées en sous-tâches (HIT) et leur exécution est **payée**.



[Your Account](#)[HITS](#)[Qualifications](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Amazon Mechanical Turk
Artificial Intelligence

Already have an account
Sign in as a [Worker](#) | [Requester](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Mechanical Turk is a marketplace for work.
We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.
179,373 HITs available. [View them now.](#)

Make Money
by working on HITs

HITs - Human Intelligence Tasks - are individual tasks that you work on. [Find HITs now.](#)

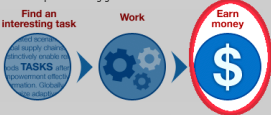
As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task

Work

Earn money



[Find HITs Now](#)

Get Results
from Mechanical Turk Workers

Ask workers to complete HITs - Human Intelligence Tasks - and get results using Mechanical Turk. [Register Now](#)


As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Fund your account

Load your tasks


Get results



[Get Started](#)

Amazon Mechanical Turk

MTurk est une plate-forme de **myriadisation** du travail parcellisé : les tâches sont découpées en sous-tâches (HIT) et leur exécution est **payée**.



[Your Account](#)[HITS](#)[Qualifications](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Amazon Mechanical Turk
Artificial Intelligence

Already have an account?
[Sign in as a Worker](#) | [Requester](#)

Mechanical Turk is a marketplace for work.

We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient.

179,373 HITS available. [View them now.](#)

Make Money by working on HITS

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)


As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task

Work

Earn money



[Find HITs Now](#)

Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Register Now](#)


As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- **Get thousands of HITs completed in minutes**
- **Pay only when you're satisfied with the results**

Fund your account

Load your tasks

Get results



[Get Started](#)

Caractéristiques d'AMT

Rémunération :

- ▶ à la tâche (*illégal* en France sauf (rares) exceptions) : moins de 2 \$/h
- ▶ pas de relation explicite entre les *Turkers* et les *Requesters*

Tâches :

- ▶ nouveaux usages : par exemple, des créations artistiques, comme <http://www.thesheepmarket.com/>
- ▶ des tâches traditionnellement réalisées par des employés salariés : transcription, traduction (agences LDC, ELDA), etc

AMT : le rêve devenu réalité ?

Cheap and Fast — But is it Good? **Evaluating Non-Expert Annotations for Natural Language Tasks**

Rion Snow[†] Brendan O'Connor[‡] Daniel Jurafsky[§] Andrew Y. Ng[†]

[†]Computer Science Dept.
Stanford University
Stanford, CA 94305

{rion,ang}@cs.stanford.edu

[‡]Dolores Labs, Inc.
832 Capp St.
San Francisco, CA 94110

brendano@doloreslabs.com

[§]Linguistics Dept.
Stanford University
Stanford, CA 94305

jurafsky@stanford.edu

[Snow et al., 2008]

AMT : le rêve devenu réalité ?

Cheap and Fast — But is it Good? **Evaluating Non-Expert Annotations for Natural Language Tasks**

Rion Snow[†] Brendan O'Connor[‡] Daniel Jurafsky[§] Andrew Y. Ng[†]

[†]Computer Science Dept.
Stanford University
Stanford, CA 94305
`{rion,ang}@cs.stanford.edu`

[‡]Dolores Labs, Inc.
832 Capp St.
San Francisco, CA 94110
`brendano@doloreslabs.com`

[§]Linguistics Dept.
Stanford University
Stanford, CA 94305
`jurafsky@stanford.edu`

[Snow et al., 2008]

C'est **très peu cher, rapide, de bonne qualité**
et c'est un **hobby** pour les *Turkers* !

AMT permet de réduire les coûts

Très basse rémunération \Rightarrow coûts faibles ? Oui, mais. . .

- ▶ coût de mise au point de l'**interface**
- ▶ coût de création de protections contre les **spammers**
- ▶ coût de **validation** et de **post-traitement**

certaines tâches (par exemple, la traduction du pachto vers l'anglais) génèrent des coûts similaires aux coûts habituels dans le domaine, du fait du **manque de Turkers qualifiés** [Novotney and Callison-Burch, 2010].

AMT permet de produire des ressources de qualité ?

- ▶ permet de produire des ressources de qualité dans certains cas précis (par exemple, la transcription simple)
- ▶ mais :
 - ▶ la qualité est insuffisante lorsque la tâche est **complexe** (par exemple, le résumé [Gillick and Liu, 2010])
 - ▶ l'**interface** d'AMT pose parfois problème [Tratz and Hovy, 2010]
 - ▶ les *Turkers* posent parfois problème (tricheurs, **spammers**)
 - ▶ le modèle de rémunération **à la tâche** pose problème [Kochhar et al., 2010]
- ▶ pour certaines tâches simples les outils de TAL produisent de **meilleurs résultats** qu'AMT [Wais et al., 2010].

AMT : un passe-temps pour les Turkers ?

[Ross et al., 2010, Ipeirotis, 2010] montre que :

- ▶ Turkers sont avant tout motivés par l'**argent** (91 %) :
 - ▶ 20 % considèrent AMT comme leur source de revenu primaire ;
 - ▶ 50 % comme leur source de revenu secondaire ;
 - ▶ l'aspect loisir n'est important que pour une minorité (30 %).
- ▶ 20 % des Turkers passent plus de 15 h par semaine sur AMT, et contribuent à 80 % des tâches.
- ▶ le salaire horaire moyen observé est **inférieur à 2 \$**.

Est-ce qu'AMT est éthique et/ou légal ?

Éthique :

- ▶ pas d'**identification** : pas de lien officiel entre *Requesters* et *Turkers* et entre *Turkers*
- ▶ pas de possibilité de **se syndiquer**, pour protester contre des manquements des *Requesters* ou ester en justice
- ▶ pas de **salaire minimum** (< 2 \$/h en moyenne)
- ▶ possibilité de **refuser de payer** les *Turkers*

Est-ce qu'AMT est éthique et/ou légal ?

Mechanical Turk is a marketplace for work.

We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.

179,373 HITs available. [View them now.](#)

Make Money by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work



Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Register Now](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results



Est-ce qu'AMT est éthique et/ou légal ?

Légalité :

- ▶ accord de licence d'Amazon : les *Turkers* sont considérés comme des travailleurs indépendants \Rightarrow ils sont supposés se déclarer comme tels et payer les cotisations afférentes
- ▶ illusoire, vu le niveau de rémunération
- \Rightarrow les États perdent une source de revenus légitime



Bartle, R. (1996).

Hearts, clubs, diamonds, spades : Players who suit MUDs.
The Journal of Virtual Environments, 1(1).



Candito, M. and Seddah, D. (2012).

Le corpus Sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical.
In Proceedings of the Traitement Automatique des Langues Naturelles (TALN), Grenoble, France.



Chamberlain, J., Poesio, M., and Kruschwitz, U. (2008).

Phrase Detectives : a web-based collaborative annotation game.

In Proceedings of the International Conference on Semantic Systems (I-Semantics'08), Graz, Austria.



Cooper, S., Treuille, A., Barbero, J., Leaver-Fay, A., Tuite, K., Khatib, F., Snyder, A. C., Beenen, M., Salesin, D., Baker, D., and Popović, Z. (2010).

The challenge of designing scientific discovery games.

In Proceedings of the Fifth International Conference on the Foundations of Digital Games, FDG '10, pages 40–47, New York, NY, USA. ACM.



Deneme (2009).

How many turkers are there?

<http://groups.csail.mit.edu/uid/deneme/>.



Fort, K., Adda, G., and Cohen, K. B. (2011).

Amazon Mechanical Turk : Gold mine or coal mine?

Computational Linguistics (editorial), 37(2) :413–420.



Fort, K., Nazarenko, A., and Rosset, S. (2012).

Modeling the complexity of manual annotation tasks : a grid of analysis.

In International Conference on Computational Linguistics (COLING), pages 895–910, Mumbai, India.



Gillick, D. and Liu, Y. (2010).

Non-expert evaluation of summarization systems is risky.

In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk, CSLDAMT '10, pages 148–151, Stroudsburg, PA, USA. Association for Computational Linguistics.



Guillaume, B. and Perrier, G. (2015).
Dependency Parsing with Graph Rewriting.

In

Proceedings of IWPT 2015, 14th International Conference on Parsing
pages 30–39, Bilbao, Spain.



Howe, J. (2006).
The rise of crowdsourcing.
Wired Magazine, 14(6).



Ipeirotis, P. (2010).
The new demographics of mechanical turk.
[http ://behind-the-enemy-lines.blogspot.com/2010/03/new-demographics-of-mechanical-turk.html](http://behind-the-enemy-lines.blogspot.com/2010/03/new-demographics-of-mechanical-turk.html).



Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., Popović, Z., et al. (2011).

Crystal structure of a monomeric retroviral protease solved by protein folding game players.

Nature structural & molecular biology, 18(10) :1175–1177.



Kochhar, S., Mazzocchi, S., and Paritosh, P. (2010).

The anatomy of a large-scale human computation engine.

In Proceedings of Human Computation Workshop at the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2010, Washington D.C.



Lafourcade, M. and Joubert, A. (2008).

JeuxDeMots : un prototype ludique pour l'émergence de relations entre termes.

In Proceedings of the Journées internationales d'Analyse statistique des Données Textuelles (JADT), Lyon, France.



Novotney, S. and Callison-Burch, C. (2010).

Cheap, fast and good enough : automatic speech recognition with non-expert transcription.

In Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), HLT'10, pages 207–215, Stroudsburg, PA, USA. Association for Computational Linguistics.



Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. (2010).

Who are the crowdworkers ? : shifting demographics in mechanical turk.

In Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems, CHI EA '10, pages 2863–2872, New York, NY, USA. ACM.



Snow, R., O'Connor, B., Jurafsky, D., and Ng., A. Y. (2008).

Cheap and fast - but is it good ? evaluating non-expert annotations for natural language tasks.

In Proceedings of EMNLP 2008, pages 254–263.



Tratz, S. and Hovy, E. (2010).

A taxonomy, dataset, and classifier for automatic noun compound interpretation.

In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, pages 678–687, Uppsala, Sweden. Association for Computational Linguistics.



Urieli, A. (2013).

Robust French syntax analysis : reconciling statistical methods and linguistic knowledge in the Talismane toolkit.

PhD thesis, Université de Toulouse II le Mirail, France.



Wais, P., Lingamneni, S., Cook, D., Fennell, J., Goldenberg, B., Lubarov, D., Marin, D., and Simons, H. (2010).

Towards building a high-quality workforce with mechanical turk.

In Proceedings of Computational Social Science and the Wisdom of Crowds (NIPS).