

Sélectionné.e par une IA ?

Algorithmes, inégalités, et les « humains dans la boucle »

Paola Tubaro, LRI-CNRS

Centre D'Alembert, 18 avril 2019

SÉLECTION ET « BIAIS ALGORITHMIQUE »

Sélectionner les étudiants en médecine

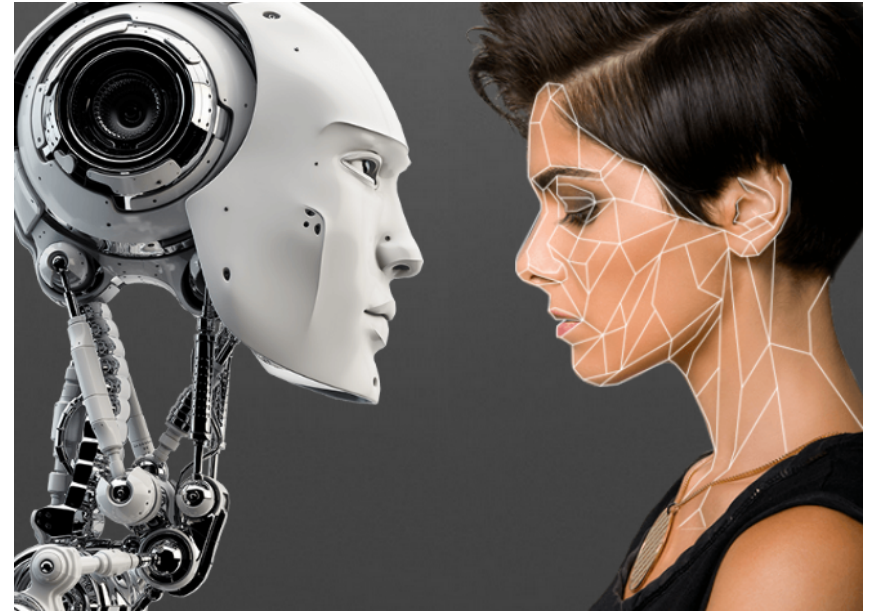


- Algorithme écrit dans les années 1980 pour alléger la tâche des recruteurs (~2500 candidatures / an)
 - 90-95% similarité avec les choix des humains
 - Mais biais systématique contre les femmes et les minorités ethniques
- ⇒ L'algorithme reproduisait et explicitait les biais humains

**DES BIAIS DE PLUS EN PLUS
COMPLEXES**

Et aujourd'hui?

- Beauty.AI 2016, le premier concours de beauté jugé par des machines: presque tous les gagnants ont la peau claire



Age group 18—29

Women



Lu Sophia

Age: 18
Real age prediction: 13



Margeri Ottis

Age: 27
Real age prediction: 23



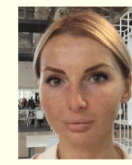
Kerri Kinney

Age: 26
Real age prediction: 18



Margarita Shestakova

Age: 25
Real age prediction: 18



Evgeniya Miruk

Age: 29
Real age prediction: 23

Des cas plus troublants...

Les algorithmes de police prédictive USA noteraient différemment les blancs et les noirs (ProPublica 2016 – mais contesté, Kamyshev 2019)

Prediction Fails Differently for Black Defendants

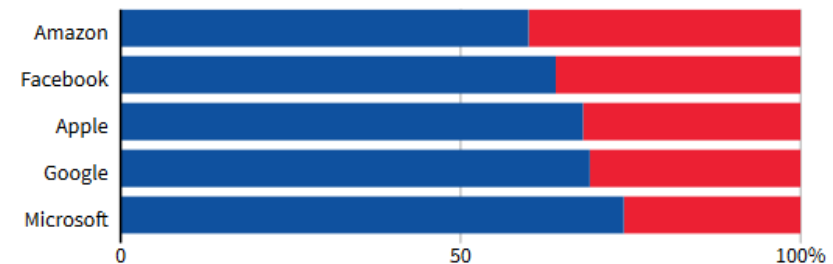
	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

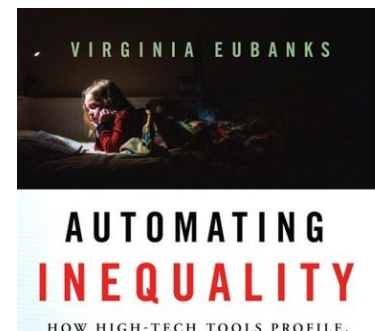
L'algorithme de recrutement de Amazon qui pénalisait les femmes, celui de Google qui leur proposait des annonces de postes moins bien payés (Datta et al 2015)

GLOBAL HEADCOUNT

■ Male ■ Female



Des décisions d'exclusion du système des aides sociales au désavantage des plus pauvres aux USA (Eubanks 2018)



Pourquoi c'est complexe

- Algorithmes pas programmés pour discriminer
- Biais réside dans le couplage algorithme-données:
 - Insuffisance ou faible qualité des données d'entraînement
 - Photos du concours de beauté incluait peu de non-blancs
 - Processus générateur des données lui-même inégalitaire
 - Les algorithmes reproduisent des biais historiques
 - Variables confondantes ou non observées

Le problème de la « black box »

- Opacité des algorithmes notamment de *deep learning*
 - Pourquoi font-ils ce qu'ils font?
 - ces technologies apprennent sans comprendre donc décident sans comprendre (Devillers 2017)
- Intérêts commerciaux et redistribution inégalitaire des profits (« Black box society », Pasquale 2015)

Un contexte difficile

- L'IA n'est pas toujours plus biaisée que des évaluateurs humains...
- ... mais elle est en train d'être intégrée très vite dans des institutions sociales clés, même si son fonctionnement reste mal connu
- Il n'existe pas de méthodes convenues pour évaluer les effets humains de l'IA tels qu'ils sont appliqués dans l'ensemble des systèmes sociaux

(Crawford & Whittaker 2016)

DE MULTIPLES PISTES

Que faire?

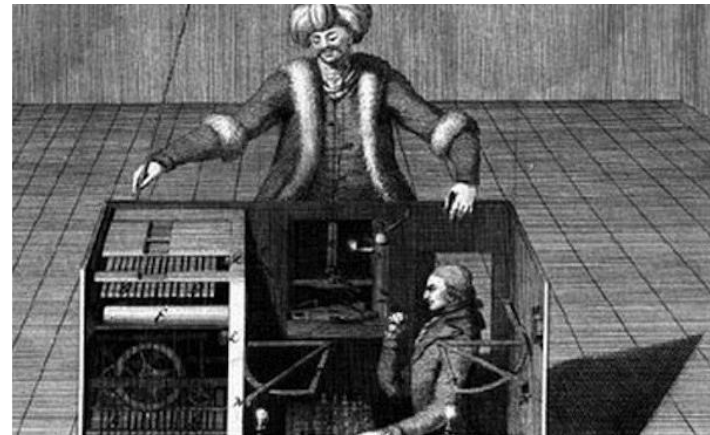
- Solutions légales: le « droit à l'explication » dans le RGPD (mais limité, Wachter et al 2017)
- Solutions algorithmiques: recherche sur
 - équité / *fairness* (Corbett-Davies & Goel 2018)
 - explicabilité, interprétabilité (Lipton 2017)
 - causalité (Diviyan)
- Action concrète pour mieux représenter les minorités (*Women in ML, Black in AI, Latinx in AI...*)
- Ouvertures inter-disciplinaires et prise en compte des conditions de production de l'IA

LA PRODUCTION DE L'IA ET L'HUMAIN-DANS-LA-BOUCLE

Tout commence par une imposture...

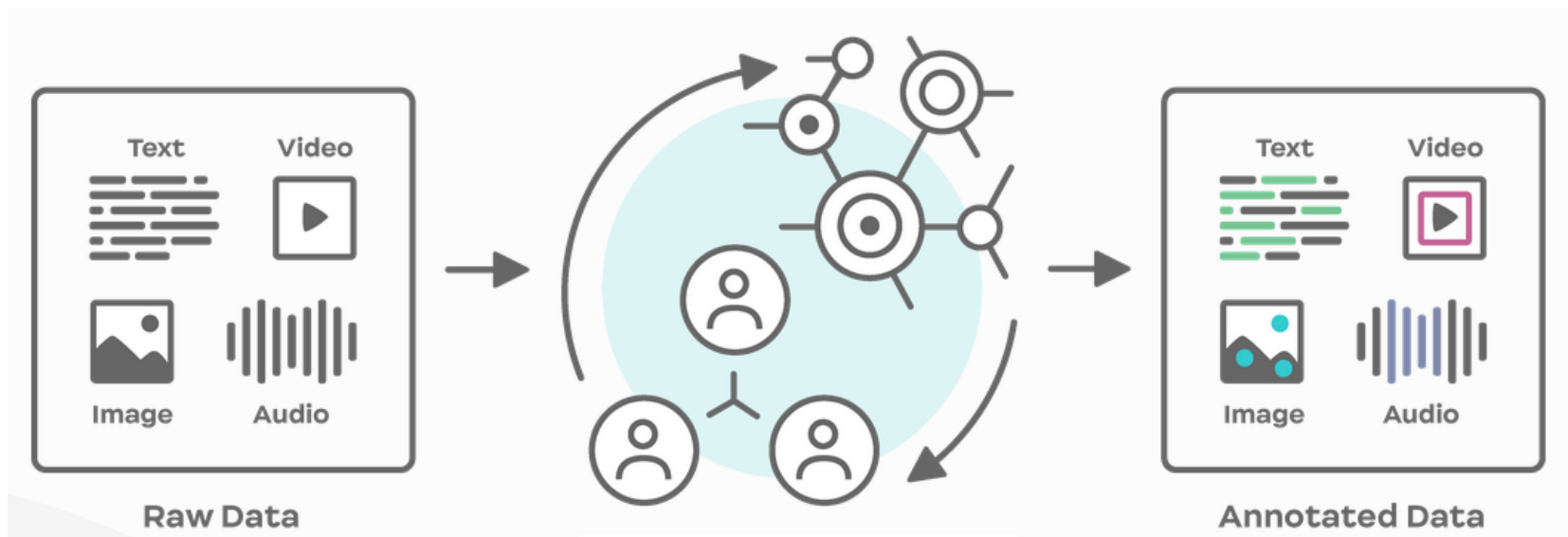
MARCH 5, 2019
FINANCIAL TIMES
Europe's AI start-ups often do not use AI, study finds

“J’ai fait mon stage dans une start-up qui promet de l’IA, mais n’en fait pas. Ils vendent des rêves et mentent sur la technologie” (Simon, 10 sept. 2017)



amazon mechanicalturk™
Artificial Artificial Intelligence

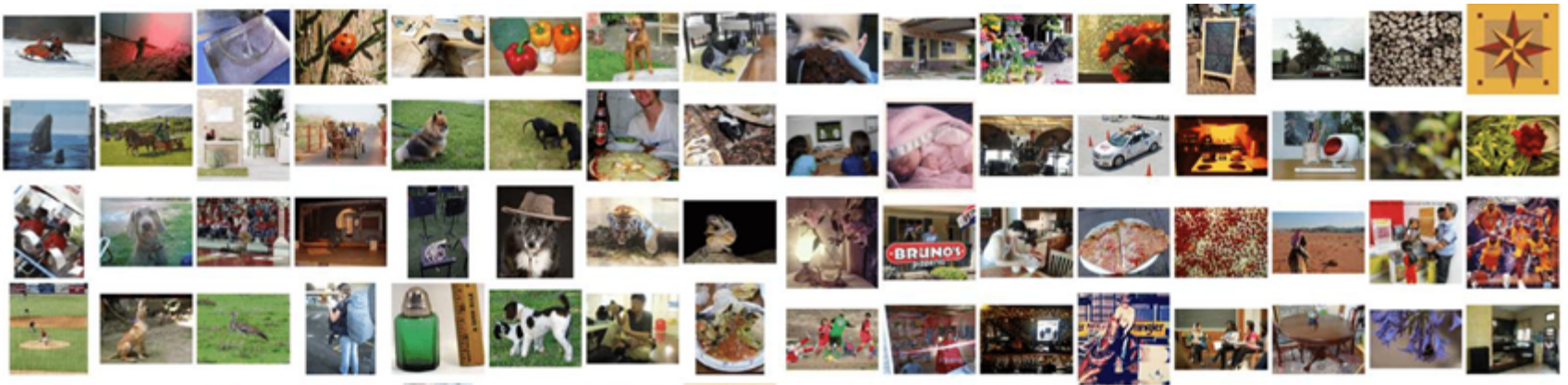
... mais il y a des vrais besoins



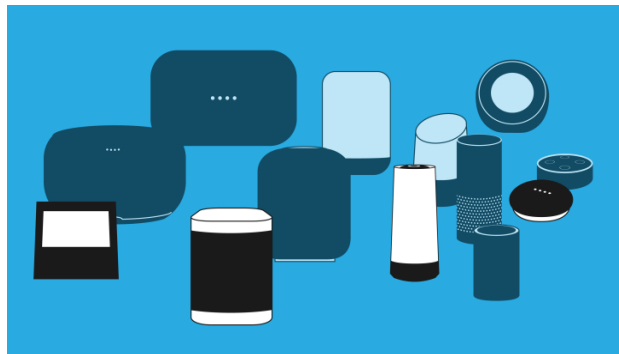
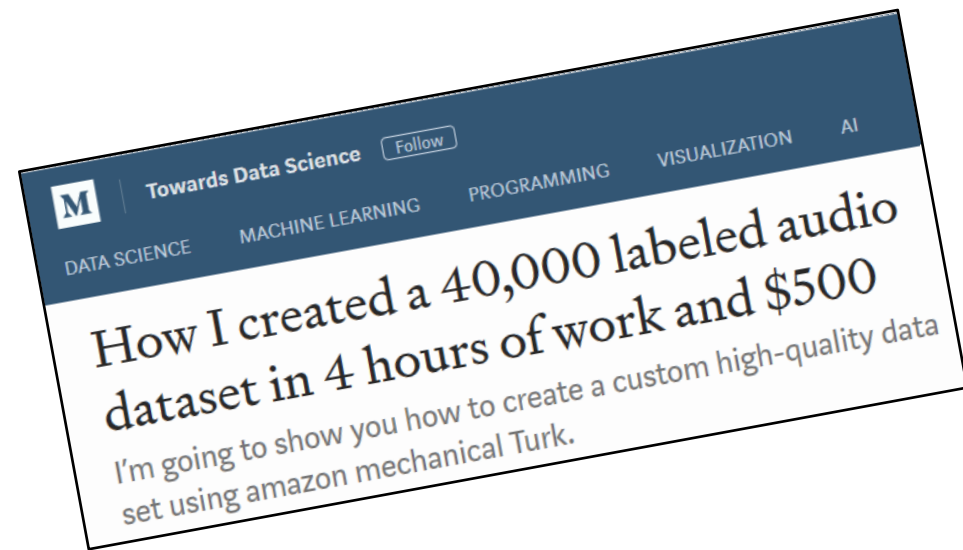
Source: Figure Eight, <https://www.figure-eight.com/>

Un exemple d'annotation: ImageNet

- 14M images annotées à la main par 50K 'Turkers' en deux ans
- Une ressource clé pour le développement de l'IA depuis 2011-12



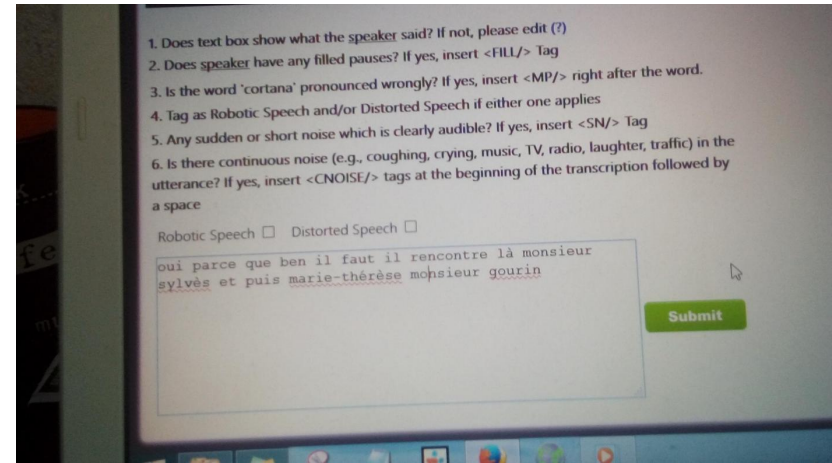
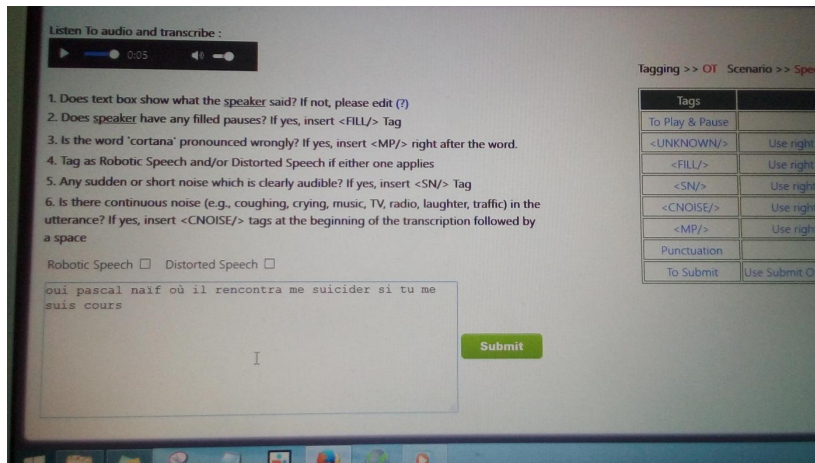
Des humains pour la génération des données



“Ma dernière tâche ? Lire à voix haute des phrases en français et m’enregistrer. C’était pour un assistant vocal.” (Kévin, 19 mai 2018)

Des besoins à tous les stades de la production: contrôles de qualité

“J'écoutais un par un les enregistrements. Les pistes étaient généralement très courtes, entre 3 et 15 secondes. Ensuite un texte s'affichait, montrant ce que l'assistante vocale avait compris et retranscrit. Mon travail était de vérifier si elle avait bien compris – si ce n'était pas le cas, je devais corriger le texte, la moindre faute de compréhension, de conjugaison ou d'orthographe.”
(Juliette, 18 mai 2018)

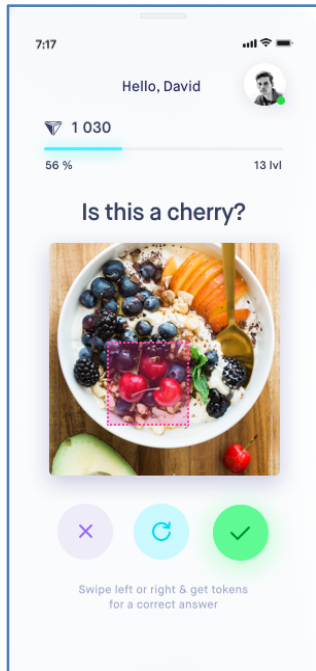


Des sources de biais

- Qui décide des labels apposés aux images ?
- Quelles voix vont être enregistrées ?
- Les conversations de qui sont écoutées ?

COMMENT L'HUMAIN FAIT PARTIE DE LA BOUCLE

L'apport humain est du 'micro-travail'



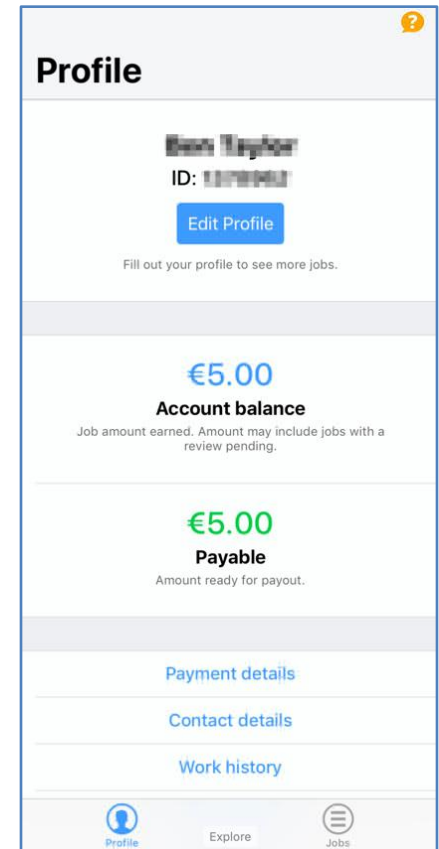
Des petites tâches fragmentées...



...réalisées en ligne, à distance...

HITS	Reward	Created	Actions	
1	\$0.30	33s ago	Preview	Accept & Work
1	\$0.90	37s ago	Preview	Accept & Work
23	\$0.30	44s ago	Preview	Accept & Work
32	\$2.20	1m ago	Preview	Accept & Work

... payées à la pièce...



... et peu rémunérées

Le modèle de la « plateforme »



The screenshot shows the top section of the microWorkers website. At the top left is the logo "microWorkers" with the tagline "work & earn or offer a micro job". To the right are links for "Blog", "API", "LOGIN", and "REGISTER". The main content area features a grid of diverse human faces. Overlaid on this grid is the text "Crowdsource your Micro Jobs" in a large white font, followed by "to more than 1,248,979 Workers worldwide, Completing 37,561,168 Tasks" in a smaller white font.

 **foulefactory**

50 000 personnes en France, mobilisables à la demande, mettent à disposition leur temps et leurs compétences pour réaliser « ensemble » vos projets.

 100% de contributeurs français

 Coûts compétitifs et maîtrisés

 Maîtrise de la qualité

 Éthique et social

 Quasi temps réel

 Réactivité immédiate



Le paysage en France



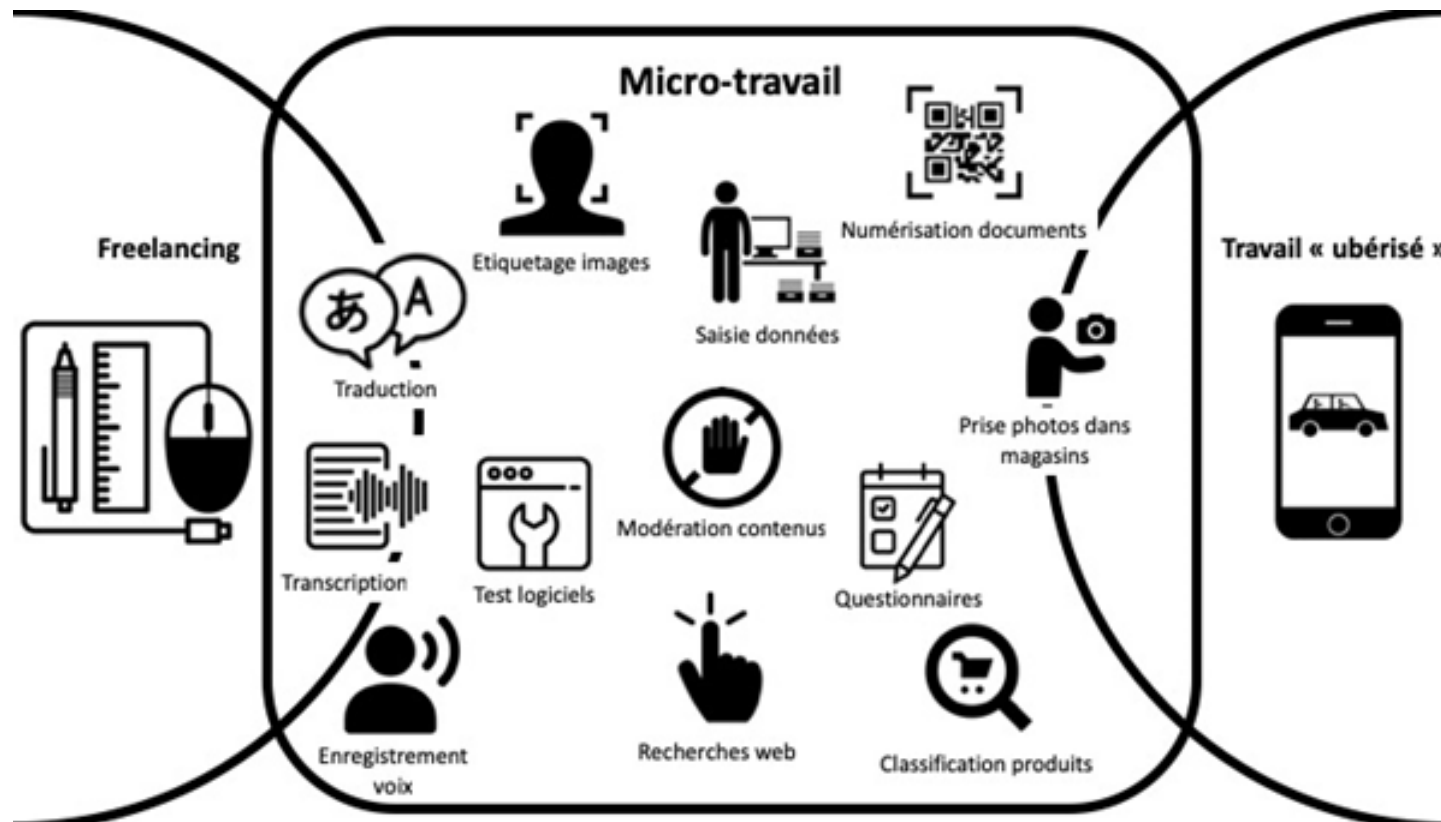
La « foule »

Step 2: verifying the images by humans

- # of synsets: 40,000 (subject to: imageability analysis)
- # of candidate images to label per synset: 10,000
- # of people needed to verify: 2-5
- Speed of human labeling: 2 images/sec (one fixation: ~200msec)
- Massive parallelism ($N \sim 10^2-3$)

$$40,000 \times 10,000 \times 3 / 2 = 600,000,000 \text{ sec} \approx \frac{19 \text{ years}}{N}$$

Le micro-travail, entre *freelancing* et *gig economy*



Une flexibilité à double tranchant

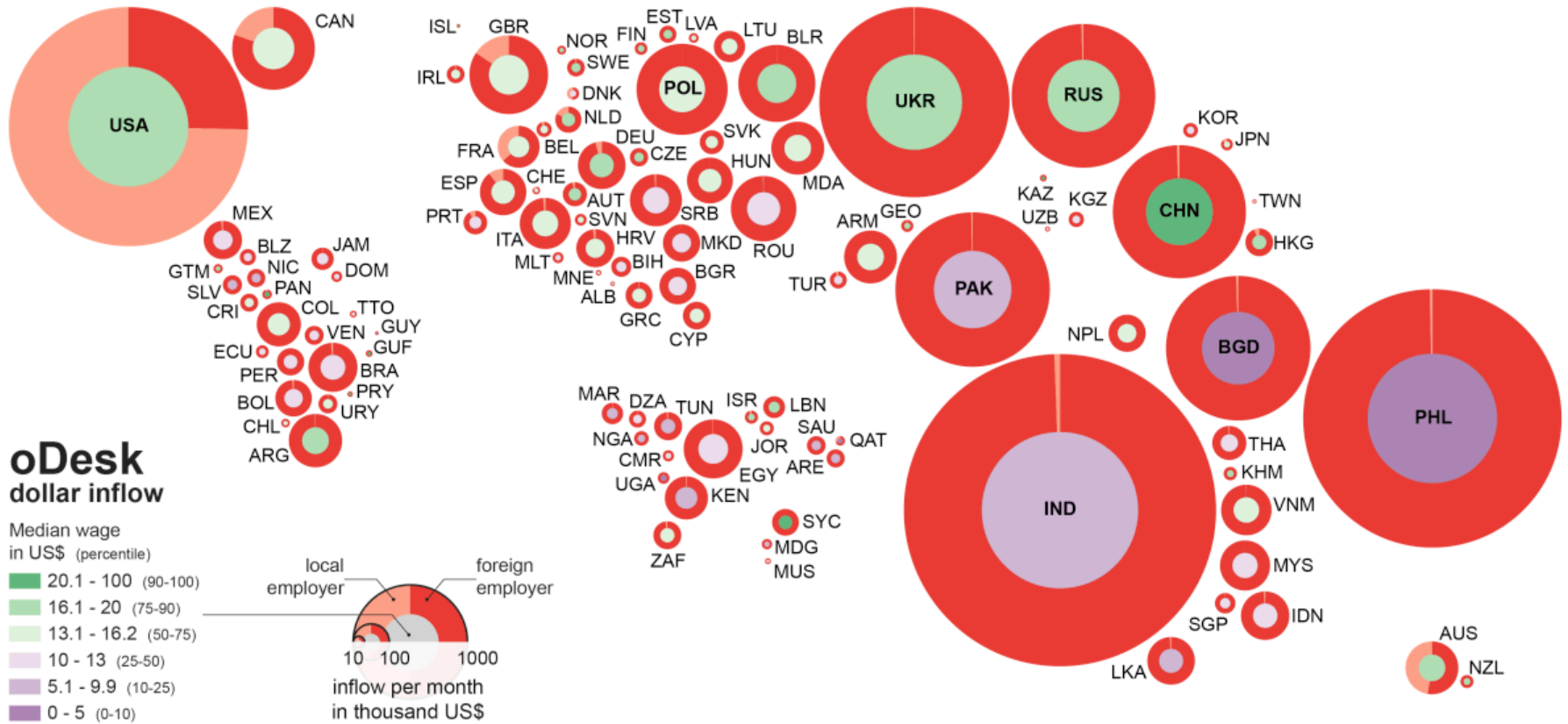
“Before the Internet, it would be really difficult to find someone, sit them down for ten minutes and get them to work for you, and then fire them after those ten minutes. But with technology, you can actually find them, pay them the tiny amount of money, and then get rid of them when you don’t need them anymore.”

L. Biewald, cofondateur de Figure Eight, 2010

Une autre forme de sélection?

- Horaires flexibles et faibles contraintes
 - Une opportunité de gains en conjoncture difficile
 - Mais...
 - Faibles rémunérations
 - Incertitude sur la demande (Graham & Anwar 2019)
 - Tâches rébarbatives, peu qualifiées (Irani 2015)
- ⇒ Qui micro-travaille ?
- ⇒ De (nouvelles) inégalités ?

LA DIVISION INTERNATIONALE DU (MICRO) TRAVAIL



Source: Graham et al (2017)

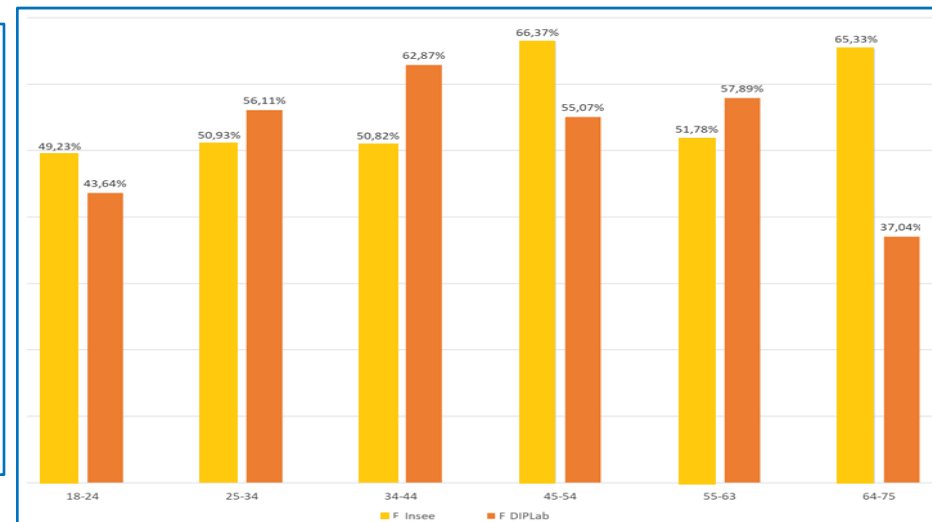
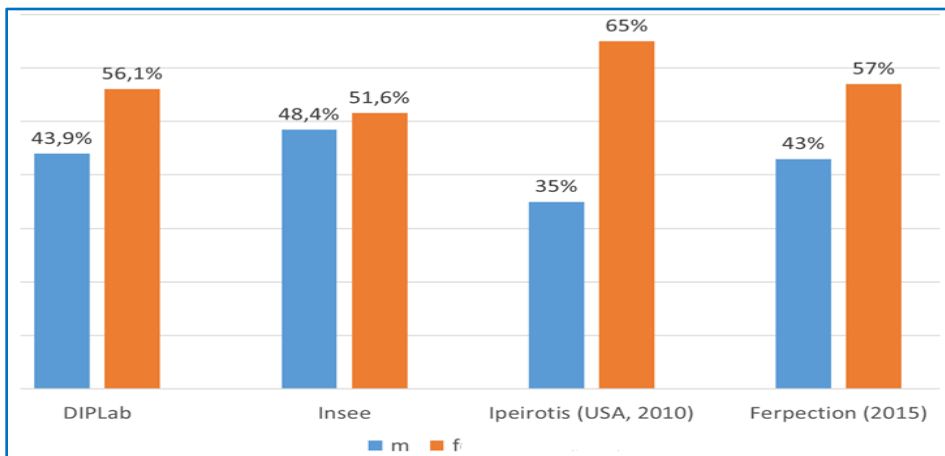
<http://diplab.eu/>

QUI SONT LES MICRO- TRAVAILLEURS EN FRANCE ?

Qui?

Surtout femmes...

... âgées 25-44



Quelles activités professionnelles?

	Micro-travail *	France †
Sans activité professionnelle	20.8%	8.9%
Etudiant.e	10.24%	3.90%
Retraité.e	5.18%	24.13%

* Source: DiPLab

† Source: Insee (2017)

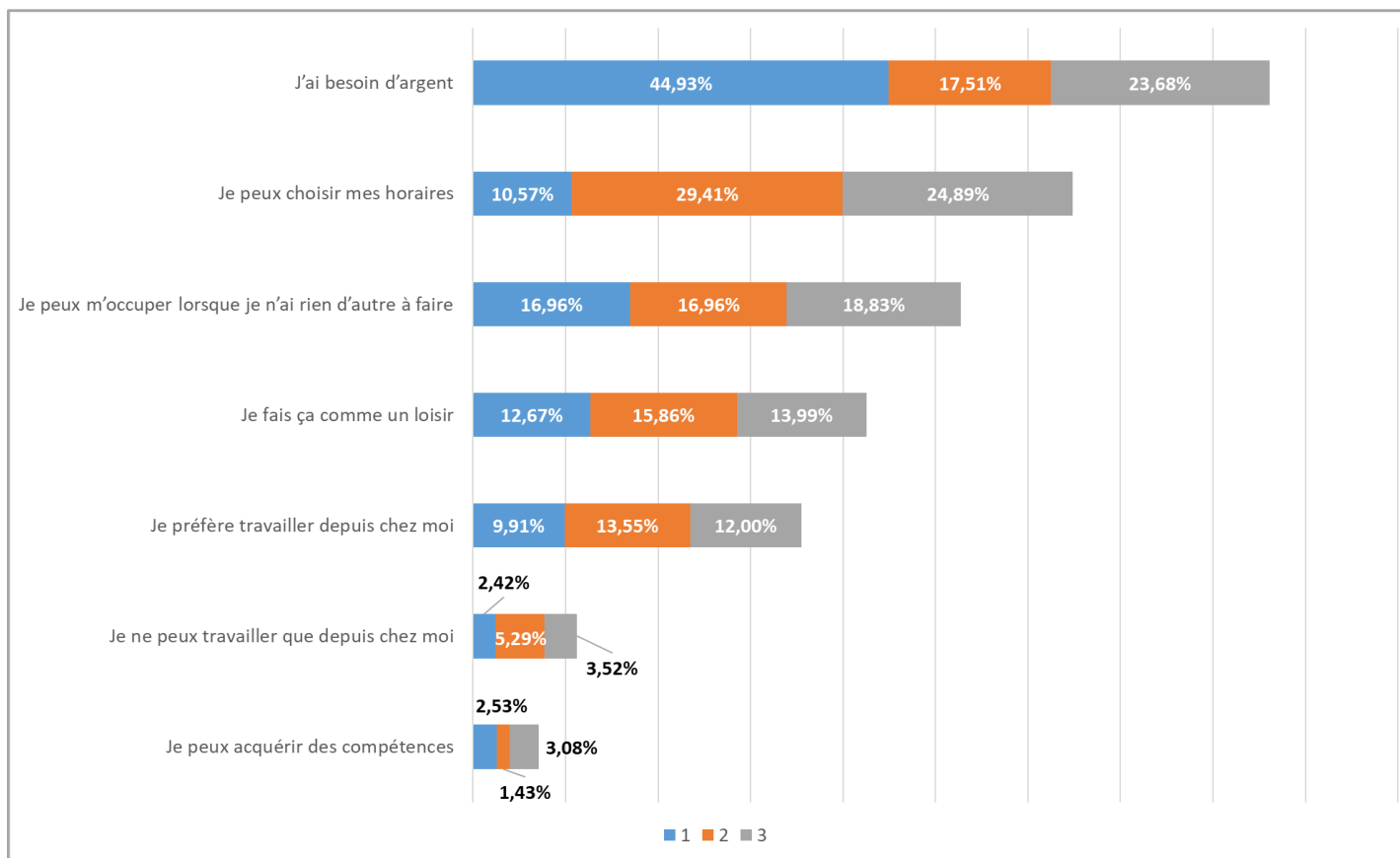
Une population à revenu faible

	Micro-travail *	France †
Classes populaires	51%	30%
<i>Sous le seuil de pauvreté (50% revenu médian)</i>	22.8%	8%
Classes moyennes	42.2%	50%
Classes supérieures	6.8%	20%

* Source: DiPLab

† Source: Insee (2018)

Pourquoi micro-travailler?



CONCLUSIONS

Conclusions

- Différentes personnes touchées par la production et le déploiement de l'IA
- La « sélection » s'opère en amont et en aval, mais n'est pas toujours reconnue comme telle
- Réflexion de plus en plus nécessaire : comment profiter des bienfaits de l'IA, de manière plus inclusive et égalitaire

Prochain RV...

MICROTRAVAIL : TECHNOLOGIE, PLATEFORMES ET EMPLOI

13/06 : CONFÉRENCE : «PERSPECTIVES POLITIQUES»
14/06 : WORKSHOP : «PERSPECTIVES DE RECHERCHE»

FRANCE STRATÉGIE
20 AVENUE DE SÉGUR, PARIS 7^E

13 / 14
2019 **JUIN**

INSCRIPTIONS OBLIGATOIRES SUR :
[HTTP://WWW.STRATEGIE.GOUV.FR](http://www.strategie.gouv.fr)

MERCI!

contact@diplab.eu

paola.tubaro@lri.fr