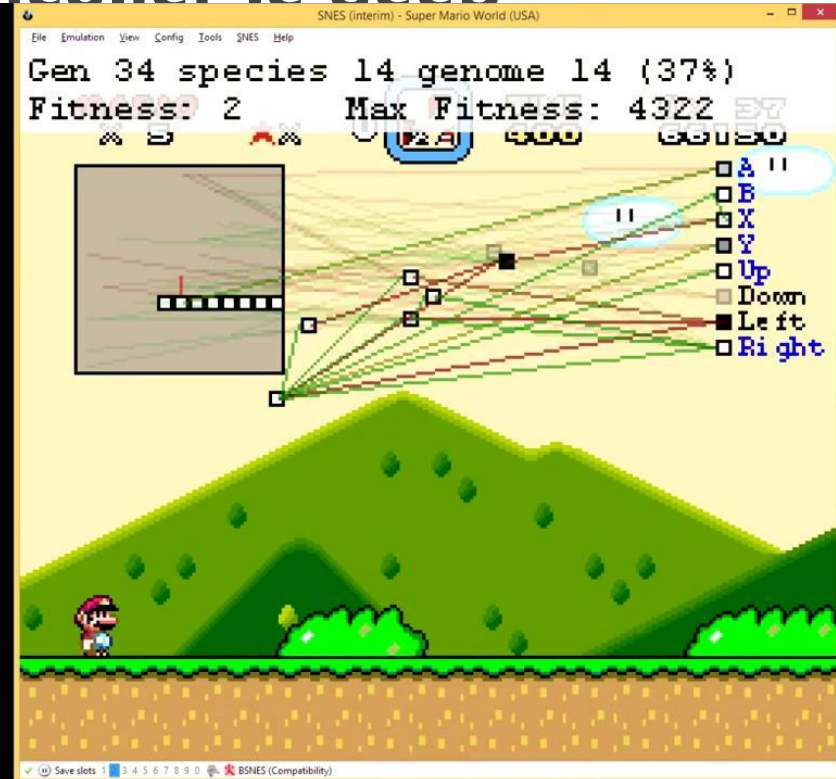
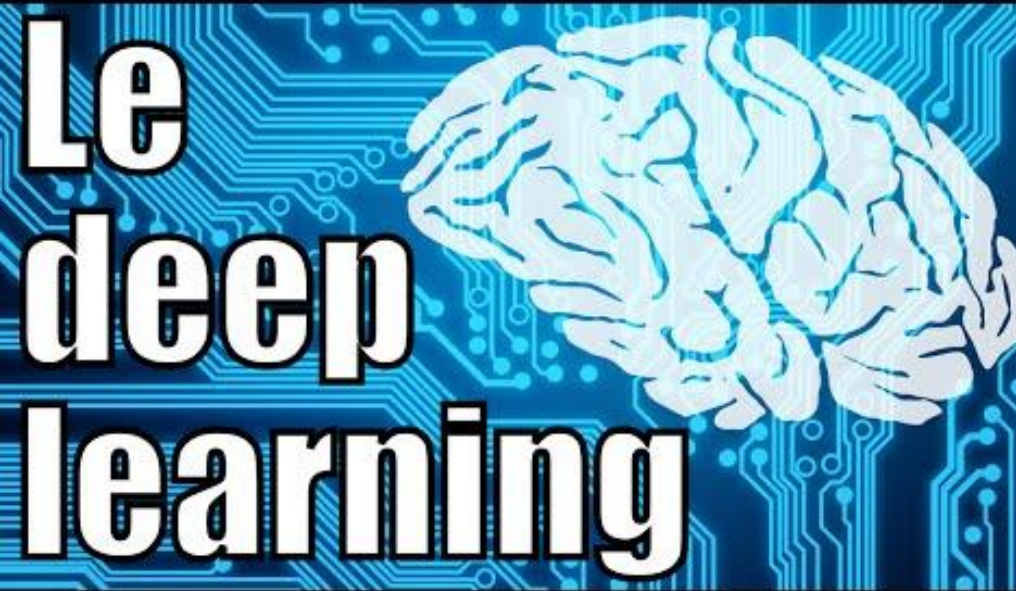




Un (tout petit) peu d'IA

Plusieurs vidéos pour expliquer ce qu'est l'IA et en particulier le deep





Comment un ordinateur comprend sans dictionnaire

“You shall know a word by the company it keeps”

(Firth, J. R. 1957:11)

C'est ce qu'on appelle la théorie **distributionnelle** du langage.

*...government debt problems turning into **banking** crises as happened in 2009...*
*...saying that Europe needs unified **banking** regulation to replace the hodgepodge...*
*...India has just given its **banking** system a shot in the arm...*

These **context words** will represent **banking**

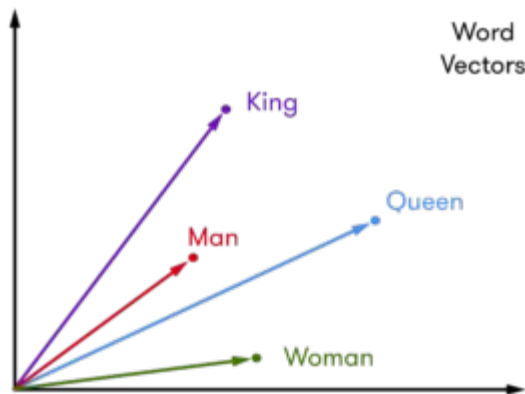


Représentation vectorielle

Rappel : un vecteur est un segment de droite **orienté**, une **direction** : on part de A pour aller vers un point B dans l'espace.

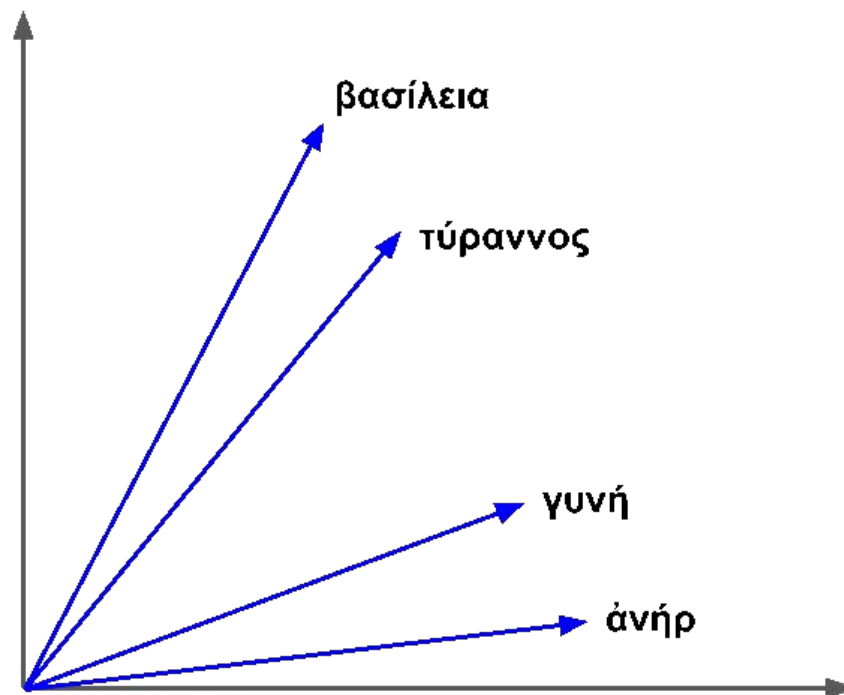
On peut avoir des vecteurs de **n -dimensions**, pas que de deux, puisque la direction elle-même peut être en n -dimensions (même si on se la représente plus facilement en 3D).

Mais on peut réduire les vecteurs à deux ou trois dimensions par commodité. Et les représenter comme ça :





Opération de vecteurs





Quelques bouts d'expériences

[les vecteurs](#)

[l'analyse de sentiments](#)

[autres exemples](#)



Les transformers



Rapide point sur les générateurs (images et paroles)

Que savez-vous sur ChatGPT ?

Que savez-vous des générateurs d'images ?

Voyons ce que peuvent faire les transformers ici :

https://github.com/OdyseusPolymetis/colabs_for_nlp/blob/main/transformers_basics.ipynb



Les transformers, ça marche comment ?

Les transformers, qu'on appelle aussi fréquemment **LLMs**, sont entraînés pour être des modèles de langue. Cela signifie qu'ils ont été formés sur de **grandes quantités de texte brut de manière auto-supervisée**.

L'apprentissage auto-supervisé est un type de formation où l'objectif est automatiquement calculé à partir des entrées du modèle.

Le modèle pré-entraîné général (souvent très gros) passe ensuite par un processus appelé **transfer learning**. Pendant ce processus, le modèle est ajusté de manière supervisée (on peut appeler ça le **finetuning**— c'est-à-dire en utilisant des étiquettes annotées par des humains — pour une tâche donnée.



L'apprentissage, un peu comme on a vu pour l'HTR

My



name

My

name



is

My

name

is



Sylvain

My

name

is

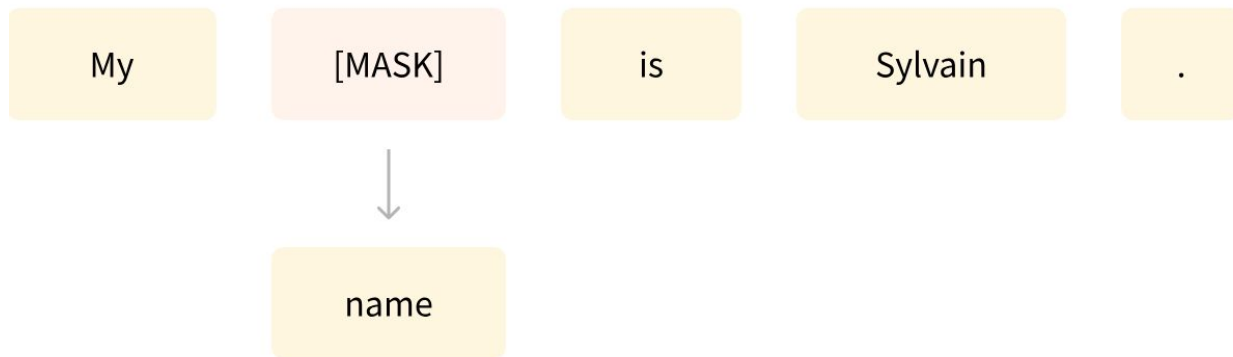
Sylvain



.

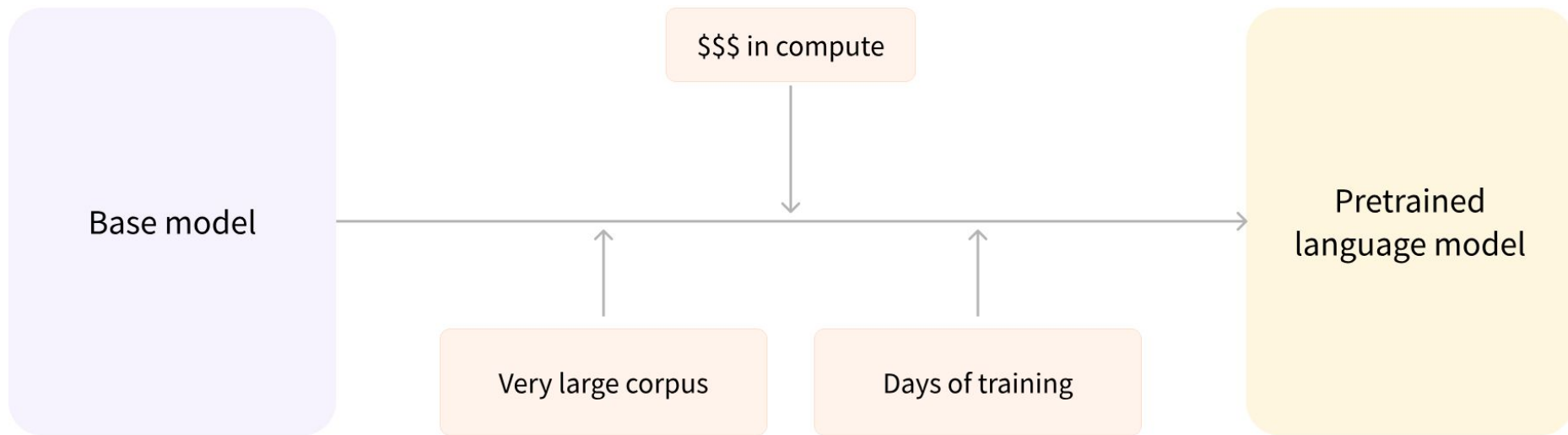


L'apprentissage, un peu comme on a vu pour l'HTR





Une chaîne très lourde



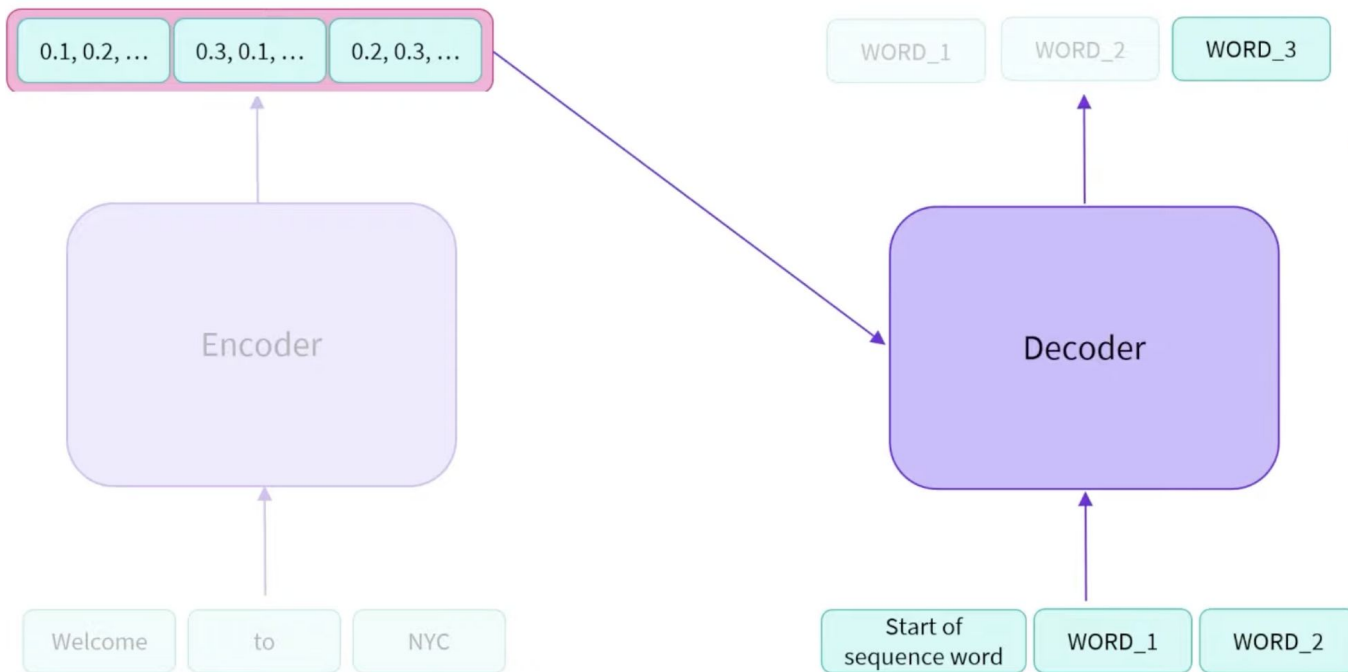


L'architecture globale



Encodeurs et décodeurs

Vidéo bien claire (mais simplifiée) sur les encodeurs : <https://youtu.be/MUqNwgPjJvQ>



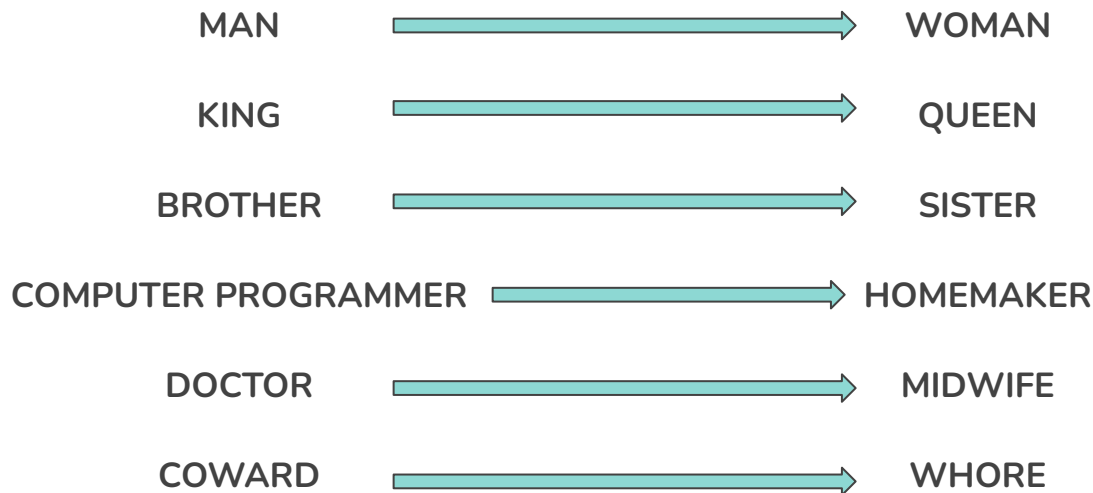


Les biais évidents





Premier exemple sur un diffuseur de vecteurs basé sur Twitter (2021)





Données triées et non triées

DeepL et GoogleTranslate pour cette expression (crédit Kaamelott) :

Français (langue détectée) ▼



Anglais (GB) ▼

Glossaire

faut m'excuser, je suis un petit peu soupe au lait en ce moment



Français ▼



Anglais ▼

faut m'excuser,
je suis un petit
peu soupe au lait
en ce moment



I have to excuse
myself, I'm a little
sluggish at the
moment.

Midjourney

a teacher



a professor



a doctor



a lawyer






Deux notebooks sur les biais de corpus

Le plus simple d'abord qu'on peut faire ensemble :


https://github.com/OdyseusPolymetis/colabs_for_nlp/blob/main/Gender_bias_in_chatgpt2.ipynb

L'autre plus long, je ne vous montrerai que les résultats mais vous pouvez tester plus tard :

https://github.com/OdyseusPolymetis/colabs_for_nlp/blob/main/biais_de_corpus.ipynb



Les biais sociaux à la loupe : 2 moteurs t2i, dalle, Stable Diffusion (SD)



Biais “ethniques”

Image
Search A



Image
Search B



DALL-E



SD





qu'est-ce qu'une "person" ?

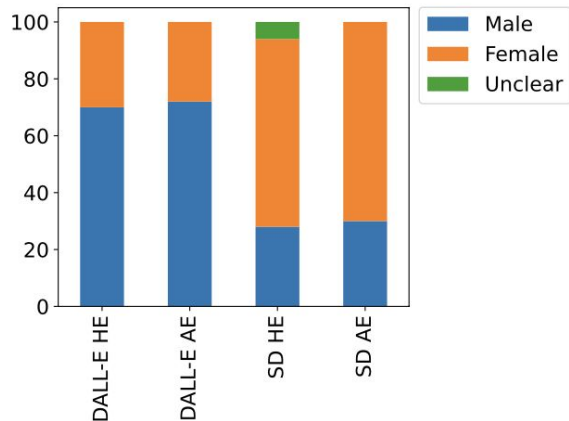


Figure 5: Gender dist. for "person"

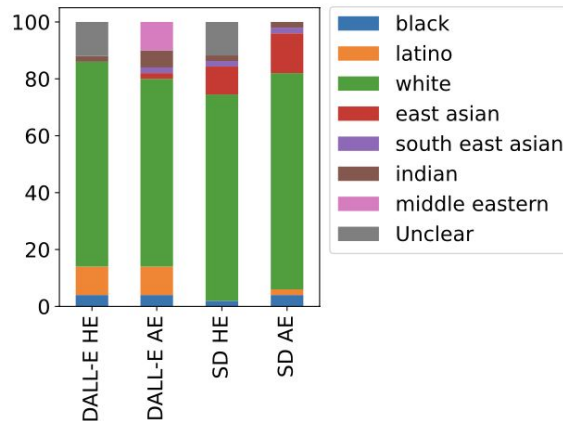


Figure 6: Race dist. for "person"

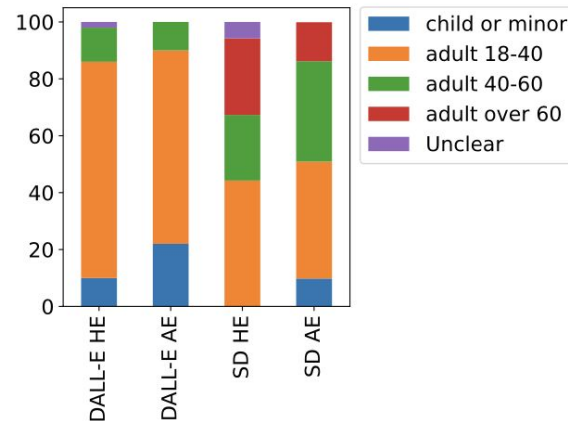
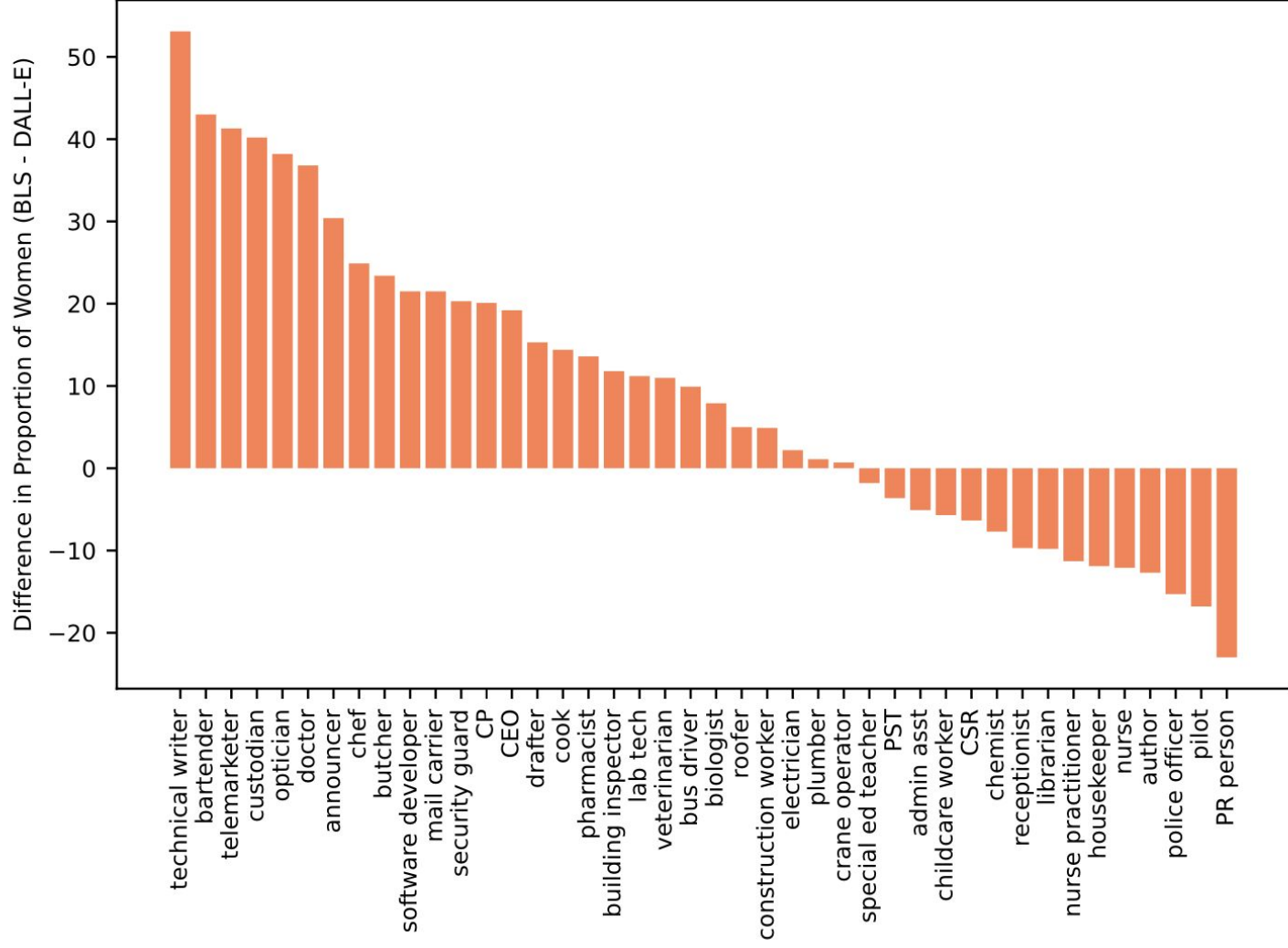


Figure 7: Age dist. for "person"

events	Nigeria	Ethiopia	PNG	India	Colombia	Mexico	Russia	Brazil	China	Australia	Germany	USA
birthday party	0.05	0.11	0.11	0.04	0.01	0.02	0.03	0.03	0.02	0.03	0.03	0.08
celebration	0.26	0.4	0.38	0.2	0.18	0.2	0.17	0.14	0.19	0.17	0.17	0.22
concert	0.16	0.18	0.22	0.06	0.05	0.05	0.01	0.02	0.05	0.04	0.03	0.03
demonstration	0.29	0.25	0.24	0.19	0.22	0.17	0.16	0.19	0.19	0.08	0.04	0.08
festival	0.24	0.25	0.22	0.09	0.08	0.06	0.04	0.03	0.03	0.13	0.09	0.07
protest	0.26	0.27	0.22	0.11	0.16	0.14	0.16	0.14	0.16	0.05	0.06	0.02
riot	0.24	0.22	0.26	0.12	0.17	0.16	0.16	0.16	0.14	0.12	0.17	0.16
wedding	0.18	0.21	0.2	0.14	0.05	0.11	0.03	0.03	0.07	0.06	0.03	0.02

events	Nigeria	Ethiopia	PNG	India	Colombia	Mexico	Russia	Brazil	China	Australia	Germany	USA
birthday party	0.31	0.28	0.25	0.28	0.07	0.06	0.08	0.05	0.16	0.04	0.04	0.02
celebration	0.44	0.31	0.34	0.32	0.34	0.22	0.3	0.29	0.26	0.17	0.29	0.25
concert	0.38	0.36	0.35	0.19	0.23	0.1	0.03	0.08	0.07	0.04	0.05	0.1
demonstration	0.38	0.33	0.33	0.36	0.31	0.25	0.26	0.31	0.27	0.15	0.23	0.13
festival	0.23	0.17	0.2	0.14	0.15	0.09	0.13	0.11	0.16	0.13	0.1	0.1
protest	0.29	0.28	0.25	0.3	0.23	0.15	0.16	0.19	0.18	0.07	0.14	0.04
riot	0.28	0.2	0.21	0.22	0.17	0.12	0.1	0.14	0.11	0.09	0.1	0.08
wedding	0.33	0.29	0.27	0.27	0.11	0.1	0.06	0.08	0.18	0.08	0.07	0.02





moyenne des représentations des traits psychologiques par ethnie (dalle et SD)

