The (Undesired) Attenuation of Human Biases by Multilinguality

Cristina España-Bonet & Alberto Barrón-Cedeño
DFKI GmbH Universitá di Bologna

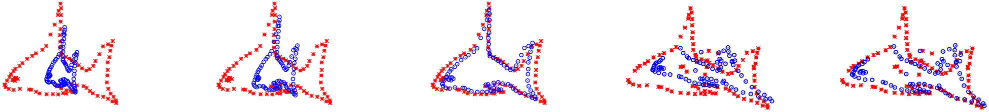
INRIA-ALMAnaCH Seminar
(extended version of EMNLP'22)

16th December 2022

Motivation

Most multilingual models *just* use a combination of monolingual corpora for training.

Are **we** distorting semantics?



[https://en.wikipedia.org/wiki/Point-set_registration]

Outline

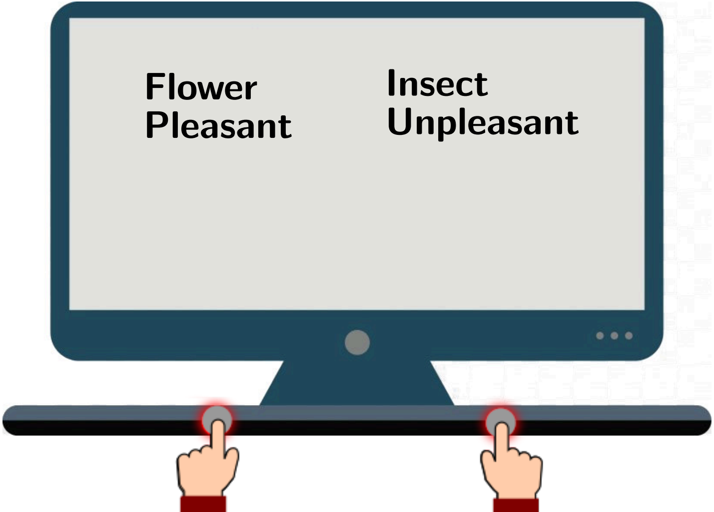
1. What is a Bias and how do we Measure them
2. Multilinguality and Cultural-Aware WEAT (CA-WEAT)
3. Experiments
4. Conclusions

IAT: Implicit Association Tests

Non-Social Human Biases

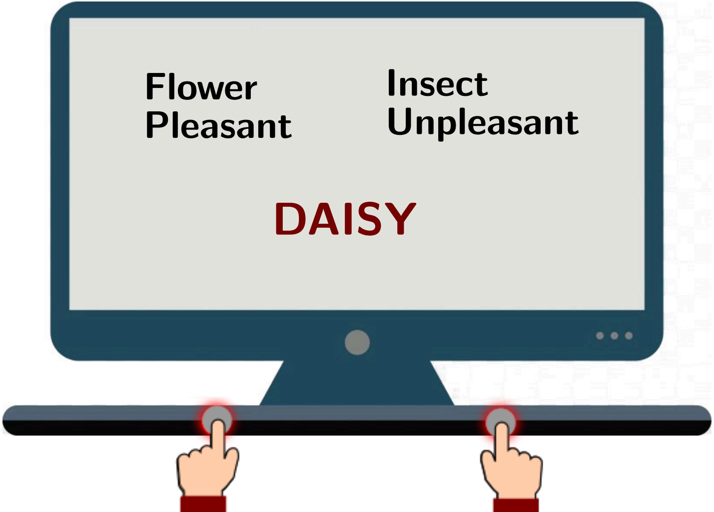
IAT: Implicit Association Tests

Non-Social Human Biases



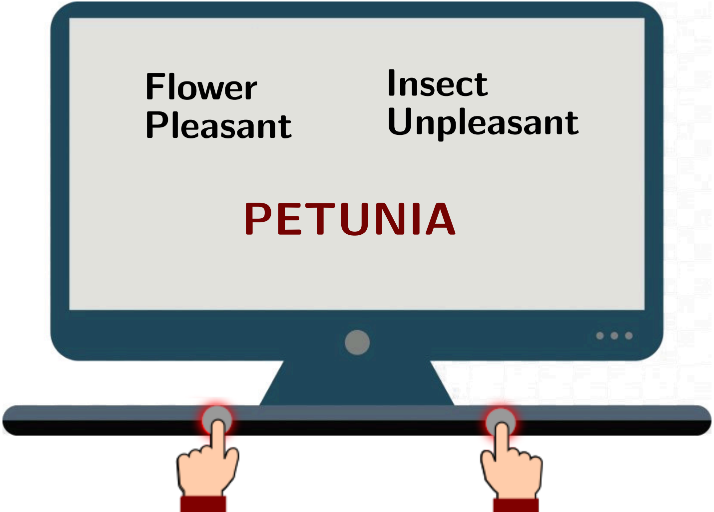
IAT: Implicit Association Tests

Non-Social Human Biases



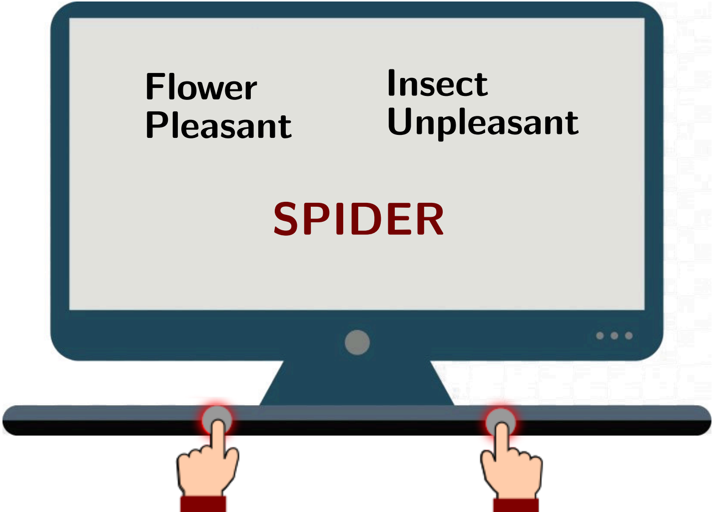
IAT: Implicit Association Tests

Non-Social Human Biases



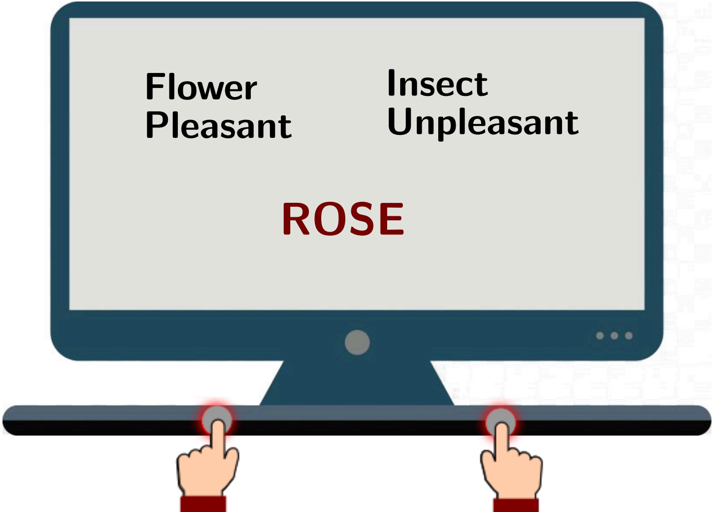
IAT: Implicit Association Tests

Non-Social Human Biases



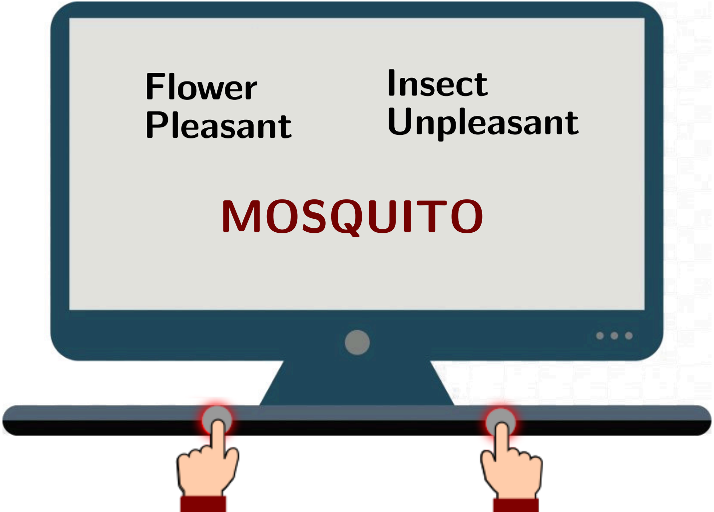
IAT: Implicit Association Tests

Non-Social Human Biases



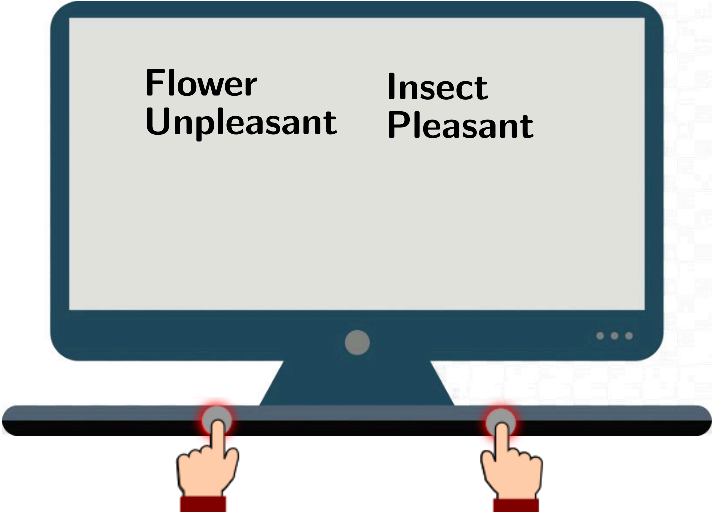
IAT: Implicit Association Tests

Non-Social Human Biases



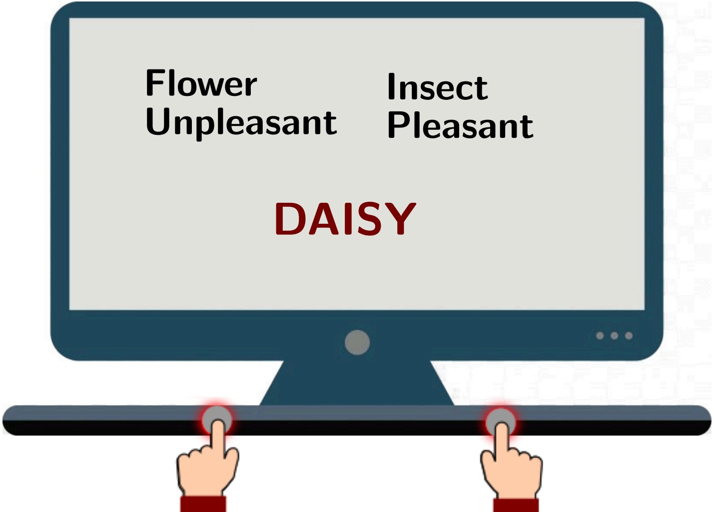
IAT: Implicit Association Tests

Non-Social Human Biases



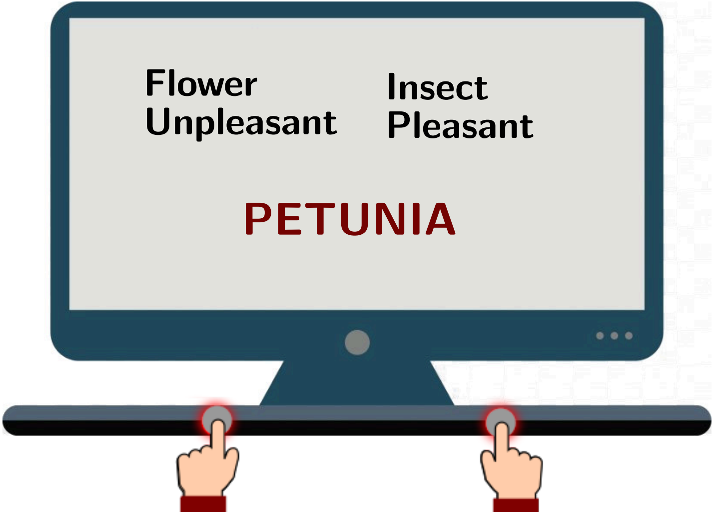
IAT: Implicit Association Tests

Non-Social Human Biases



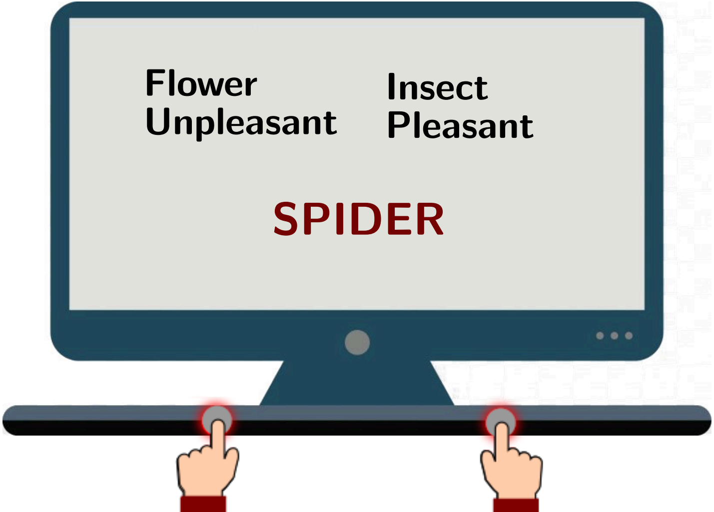
IAT: Implicit Association Tests

Non-Social Human Biases



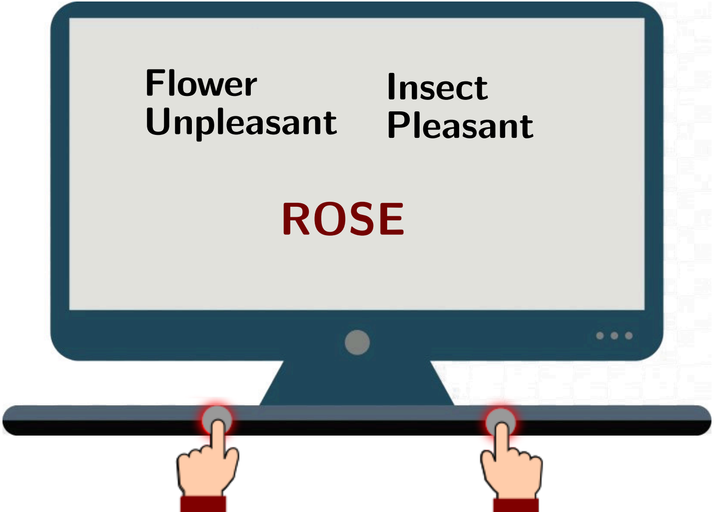
IAT: Implicit Association Tests

Non-Social Human Biases



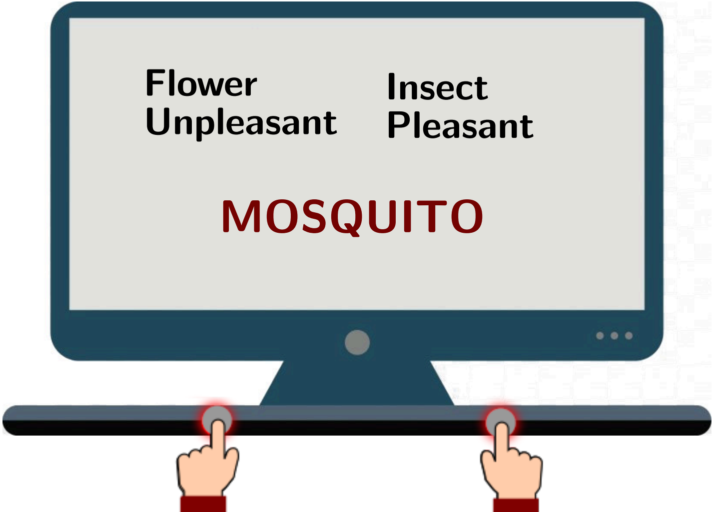
IAT: Implicit Association Tests

Non-Social Human Biases



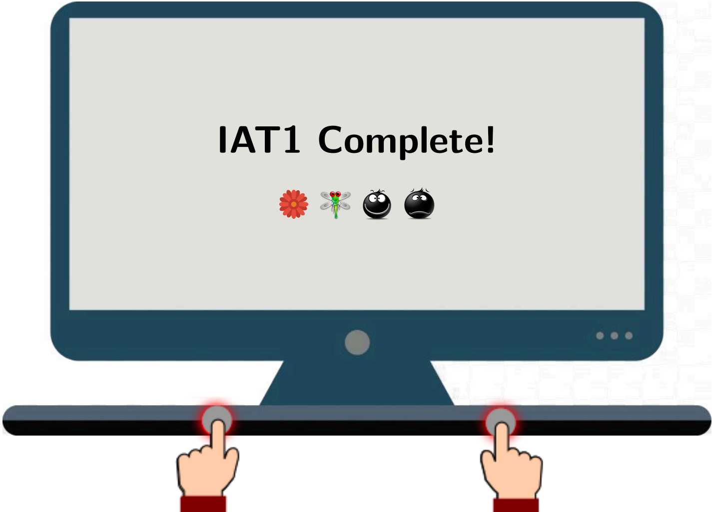
IAT: Implicit Association Tests

Non-Social Human Biases



IAT: Implicit Association Tests

Non-Social Human Biases

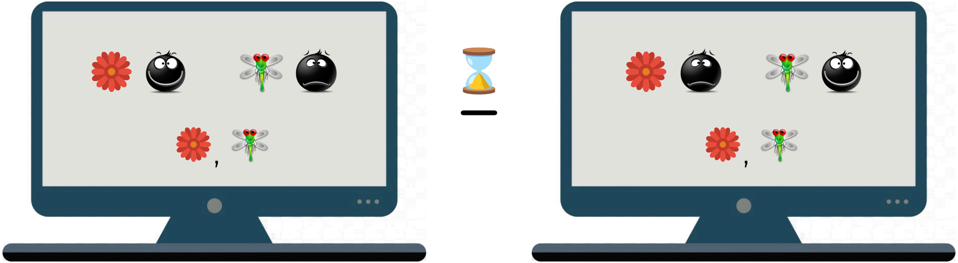


IAT: Implicit Association Tests

Non-Social Human Biases

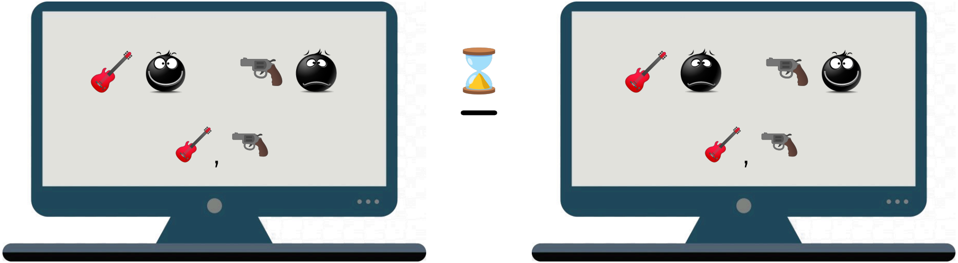
IAT1: difference in response time

(flowers & insects)



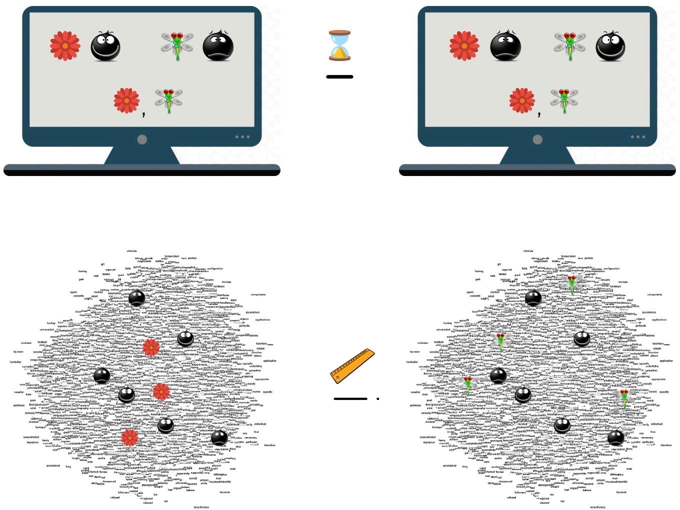
IAT2: difference in response time

(musical instruments & weapons)



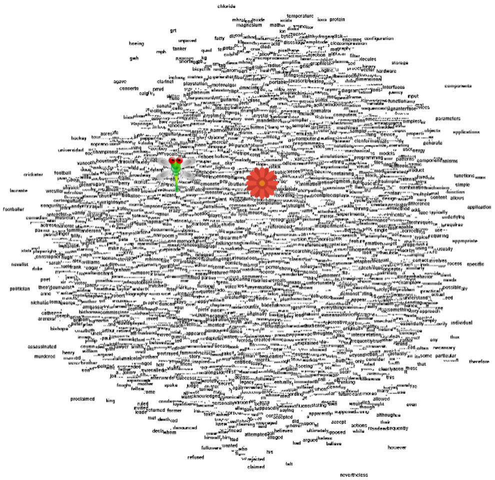
WEAT: Association Tests in Word Embeddings

WEAT, Intuition



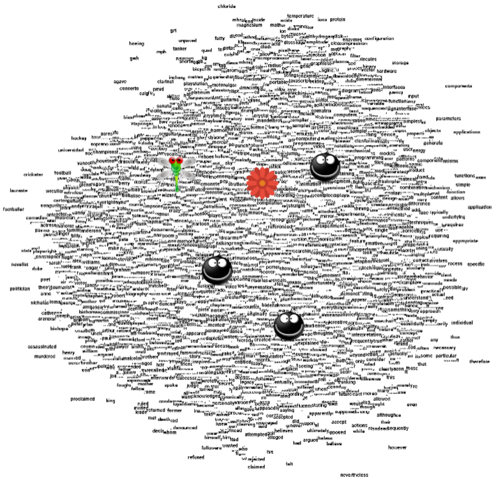
WEAT: Association Tests in Word Embeddings

Intuition, in our Embedding Space we can Measure Distances



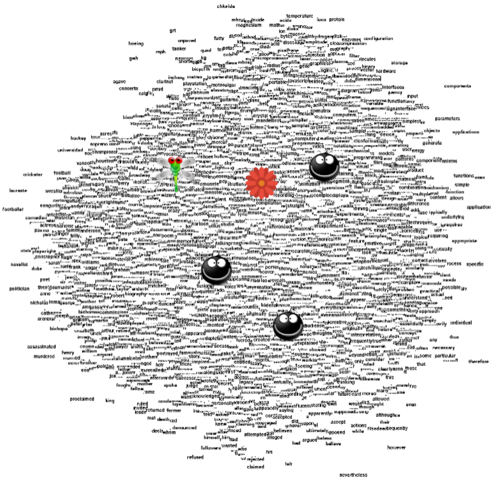
WEAT: Association Tests in Word Embeddings

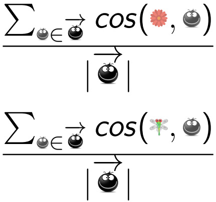
Intuition, in our Embedding Space we can Measure Distances



WEAT: Association Tests in Word Embeddings

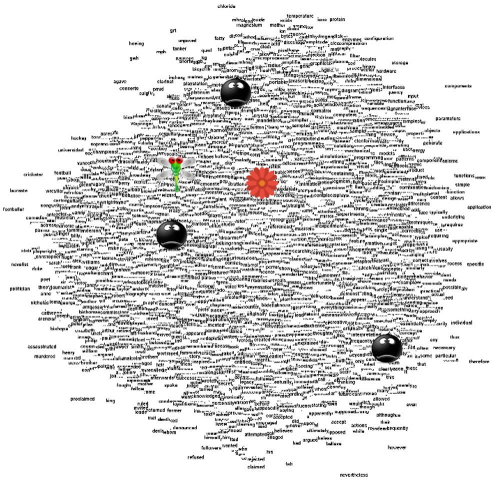
Intuition, in our Embedding Space we can Measure Distances





WEAT: Association Tests in Word Embeddings

Intuition, in our Embedding Space we can Measure Distances

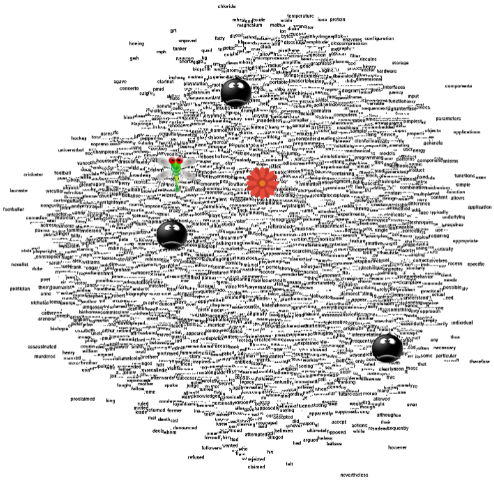


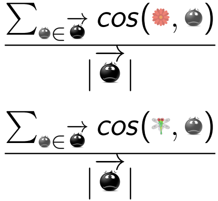
$$\frac{\sum_{\text{😟} \in \vec{\text{😟}}} cos(\text{🌼}, \text{😟})}{|\vec{\text{😟}}|}$$

$$\frac{\sum_{\text{😟} \in \vec{\text{😟}}} cos(\text{🦟}, \text{😟})}{|\vec{\text{😟}}|}$$

WEAT: Association Tests in Word Embeddings

Intuition, in our Embedding Space we can Measure Distances

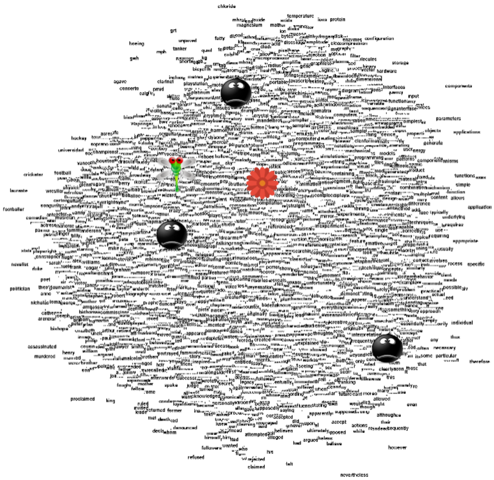




$$assoc(t, A) = \frac{\sum_{a \in A} cos(\mathbf{t}, \mathbf{a})}{|A|}$$

WEAT: Association Tests in Word Embeddings

Intuition, in our Embedding Space we can Measure Distances



$$\frac{\sum_{\text{😞} \in \vec{\text{😞}}} cos(\text{🌼}, \text{😞})}{|\vec{\text{😞}}|}$$

$$\frac{\sum_{\text{😞} \in \vec{\text{😞}}} cos(\text{🪰}, \text{😞})}{|\vec{\text{😞}}|}$$

$$assoc(t, A) = \frac{\sum_{a \in A} cos(\mathbf{t}, \mathbf{a})}{|A|}$$

$$\Delta_{assoc}(t, A, B) = assoc(t, A) - assoc(t, B)$$

WEAT: Association Tests in Word Embeddings

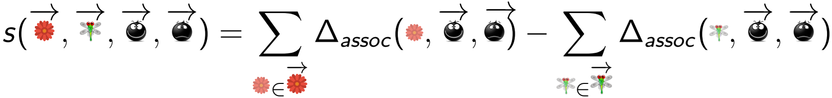
What do we Measure?

The difference in association for a term:

$$\Delta_{assoc}(t, A, B) = assoc(t, A) - assoc(t, B)$$

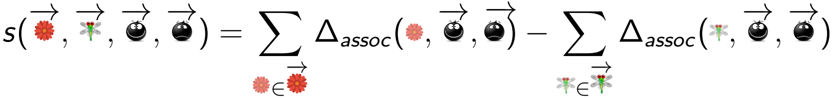
The statistic:

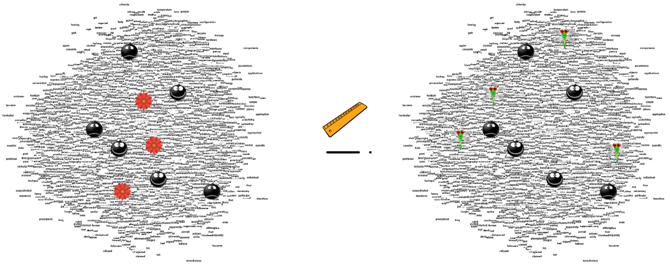
$$s(X, Y, A, B) = \sum_{x \in X} \Delta_{assoc}(x, A, B) - \sum_{y \in Y} \Delta_{assoc}(y, A, B)$$



WEAT: Association Tests in Word Embeddings

What do we Measure?





The statistic:

$$s(\vec{\text{🌺}}, \vec{\text{🦟}}, \vec{\text{😀}}, \vec{\text{😞}}) = \sum_{\text{🌺} \in \vec{\text{🌺}}} \Delta_{assoc}(\text{🌺}, \vec{\text{😀}}, \vec{\text{😞}}) - \sum_{\text{🦟} \in \vec{\text{🦟}}} \Delta_{assoc}(\text{🦟}, \vec{\text{😀}}, \vec{\text{😞}})$$

The size effect:

$$d(\vec{\text{🌺}}, \vec{\text{🦟}}, \vec{\text{😀}}, \vec{\text{😞}}) = \frac{\mu\left(\Delta_{assoc}(\text{🌺}, \vec{\text{😀}}, \vec{\text{😞}})_{\forall \text{🌺} \in \vec{\text{🌺}}}\right) - \mu\left(\Delta_{assoc}(\text{🦟}, \vec{\text{😀}}, \vec{\text{😞}})_{\forall \text{🦟} \in \vec{\text{🦟}}}\right)}{\sigma\left(\Delta_{assoc}(\text{🌺🦟}, \vec{\text{😀}}, \vec{\text{😞}})_{\forall \text{🌺🦟} \in \vec{\text{🌺}} \cup \vec{\text{🦟}}}\right)}$$

WEAT: Association Tests in Word Embeddings

Do Word Embeddings Reflect Implicit Human Associations?

[Caliskan et al., Nature, 2017]

Semantics derived automatically from language corpora contain human-like biases:

- *morally neutral* as toward insects or flowers, —our *non-social*—
- problematic as toward race or gender,
- veridical, reflecting the status quo distribution of gender with respect to careers or first names.

WEAT: Association Tests in Word Embeddings

Do Word Embeddings Reflect Implicit Human Associations?

[Caliskan et al., Nature, 2017]

Semantics derived automatically from language corpora contain human-like biases:

- *morally neutral* as toward insects or flowers, —our *non-social*—
- problematic as toward race or gender,
- veridical, reflecting the status quo distribution of gender with respect to careers or first names.

For multilinguality we need universals ⇒ **non-social only**

Outline

1. What is a Bias and how do we Measure them
   - IAT: Implicit Association Tests
   - WEAT: Association Tests in Word Embeddings
2. Multilinguality and Cultural-Aware WEAT (CA-WEAT)
3. Experiments
   - Wide Overview
   - WEAT vs X-WEAT vs CA-WEAT
   - Data Asymmetries and Isomorphism
4. Conclusions

Multilinguality and Cultural-Aware WEAT

WEAT1 and WEAT2 Original Lists

WEAT1 target items		
	Flowers	aster, clover, hyacinth, marigold, poppy, azalea, crocus, iris, orchid, rose, bluebell, daffodil, lilac, pansy, tulip, buttercup, daisy, lily, peony, violet, carnation, gladiola, magnolia, petunia, zinnia
	Insects	ant, caterpillar, flea, locust, spider, bedbug, centipede, fly, maggot, tarantula, bee, cockroach, gnat, mosquito, termite, beetle, cricket, hornet, moth, wasp, blackfly, dragonfly, horsefly, roach, weevil
WEAT2 target items		
	Instruments	bagpipe, cello, guitar, lute, trombone, banjo, clarinet, harmonica, mandolin, trumpet, bassoon, drum, harp, oboe, tuba, bell, fiddle, harpsichord, piano, viola, bongo, flute, horn, saxophone, violin
	Weapons	arrow, club, gun, missile, spear, axe, dagger, harpoon, pistol, sword, blade, dynamite, hatchet, rifle, tank, bomb, firearm, knife, shotgun, teargas, cannon, grenade, mace, slingshot, whip
WEAT1 and WEAT2 attributes		
	Pleasant	caress, freedom, health, love, peace, cheer, friend, heaven, loyal, pleasure, diamond, gentle, honest, lucky, rainbow, diploma, gift, honor, miracle, sunrise, family, happy, laughter, paradise, vacation
	Unpleasant	abuse, crash, filth, murder, sickness, accident, death, grief, poison, stink, assault, disaster, hatred, pollute, tragedy, divorce, jail, poverty, ugly, cancer, kill, rotten, vomit, agony, prison

Original version (WEAT1, WEAT2)

[Battig and Montague, 1969; Bellezza et al., 1986; Greenwald et al., 1998]

- Collected from college students in Eastern US
- Frequent terms
- Non-ambiguous terms

Multilinguality and Cultural-Aware WEAT

Original and X-WEAT Lists

Original version (WEAT1, WEAT2)

[Battig and Montague, 1969; Bellezza et al., 1986; Greenwald et al., 1998]

- Collected from college students in Eastern US
- Frequent terms
- Non-ambiguous terms

Multilingual version (X-WEAT)

[Lauscher and Glavaš, 2019; Lauscher et al., 2020]

- Literal translation
- Arabic, Croatian, German, Italian, Russian, Spanish and Turkish

Multilinguality and Cultural-Aware WEAT

Features and Issues with WEAT and X-WEAT

- **WEAT**: American English, represents the culture of the (Eastern) US
- **X-WEAT**: Multilingual, but represents the culture of the (Eastern) US!

  —and this applies to all NLP using translation—

  - duplicates? (*gun*, *pistol* $\rightarrow$ *pistolet*)
  - frequent terms? (*taon* $\rightarrow$ *horsefly*)
  - non-ambiguous terms? (*blade* $\rightarrow$ *lame*)

Multilinguality and Cultural-Aware WEAT

Features and Issues with WEAT and X-WEAT

- **WEAT**: American English, represents the culture of the (Eastern) US
- **X-WEAT**: Multilingual, but represents the culture of the (Eastern) US!

  —and this applies to all NLP using translation—

  - duplicates? (*gun*, *pistol* $\rightarrow$ *pistolet*)
  - frequent terms? (*taon* $\rightarrow$ *horsefly*)
  - non-ambiguous terms? (*blade* $\rightarrow$ *lame*)
- **CA-WEAT**: Multilingual and culturaly aware

Multilinguality and Cultural-Aware WEAT

Features and Issues with WEAT and X-WEAT (the safe version :-))

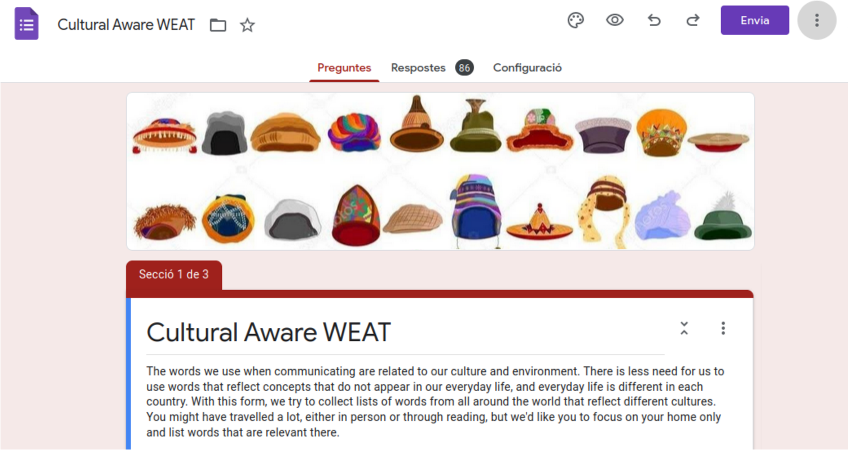
- **WEAT**: American English, represents the culture of the (Eastern) US
- **X-WEAT**: Multilingual, but represents the culture of the (Eastern) US!

  —and this applies to all NLP using translation—

  - duplicates? (*violin, fiddle* → *violín*)
  - frequent terms? (*gnat* → *jején*)
  - non-ambiguous terms? (*blade* → *hoja*)
- **CA-WEAT**: Multilingual and culturaly aware

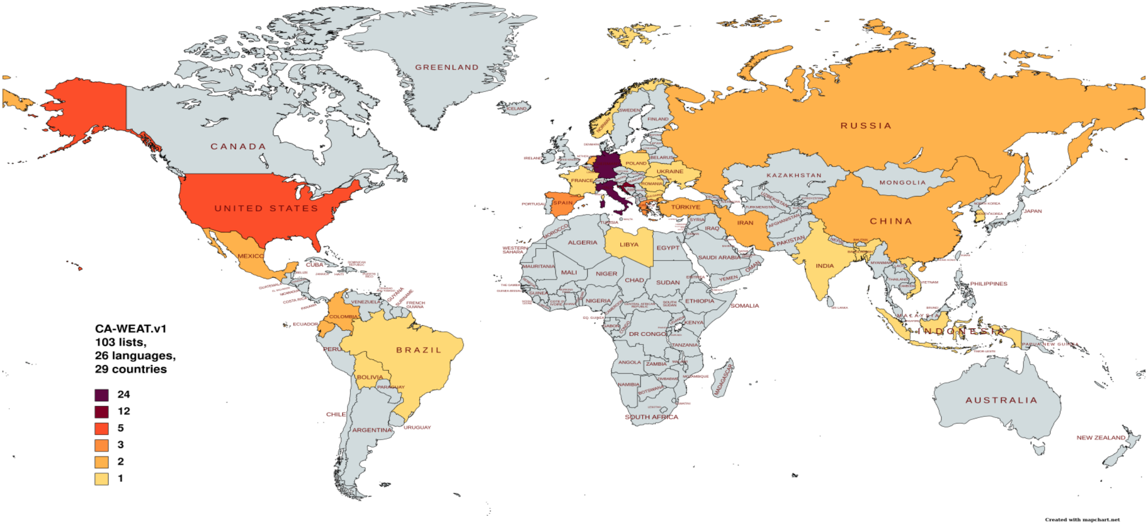
Multilinguality and Cultural-Aware WEAT

CA-WEAT



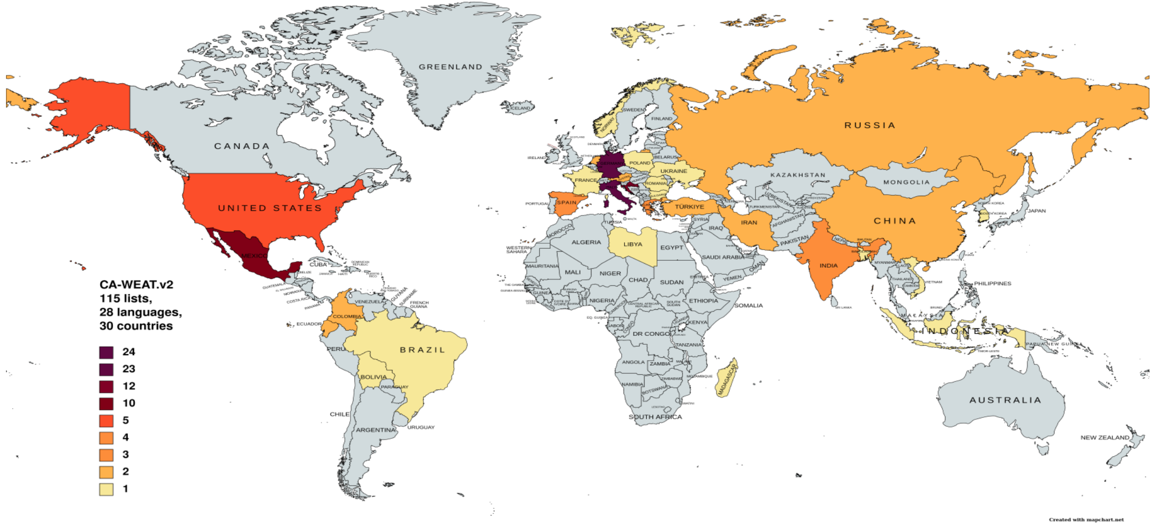
Multilinguality and Cultural-Aware WEAT

CA-WEATs per Country (not the best Distribution!)



Multilinguality and Cultural-Aware WEAT

CA-WEATs per Country (not the best Distribution!)



115 lists means 115 people.

THANKS!!

https://github.com/cristinae/CA-WEAT

Outline

1. What is a Bias and how do we Measure them
   - IAT: Implicit Association Tests
   - WEAT: Association Tests in Word Embeddings
2. Multilinguality and Cultural-Aware WEAT (CA-WEAT)
3. Experiments
   - Wide Overview
   - WEAT vs X-WEAT vs CA-WEAT
   - Data Asymmetries and Isomorphism
4. Conclusions

Experiments

Embedding Models & Languages

- **Pre-trained fastText word embeddings**

WP	WPali	CCWP

Experiments

Embedding Models & Languages

- **Pre-trained fastText word embeddings**

  WP WPali CCWP

- **Comparable word embeddings with a subset of CC-100**

  CCe CCeVMuns CCeVMsup CCe2langs CCe9langs

Embedding Models & Languages

- **Pre-trained fastText word embeddings**

  WP WPali CCWP

- **Comparable word embeddings with a subset of CC-100**

  CCe CCeVMuns CCeVMsup CCe2langs CCe9langs

- **Word embeddings extracted from contextual models at different layers**

  BERT mBERT XLM XGLM

Experiments

Embedding Models & Languages

- **Pre-trained fastText word embeddings**

  WP WPali CCWP

- **Comparable word embeddings with a subset of CC-100**

  CCe CCeVMuns CCeVMsup CCe2langs CCe9langs

- **Word embeddings extracted from contextual models at different layers**

  BERT mBERT XLM XGLM

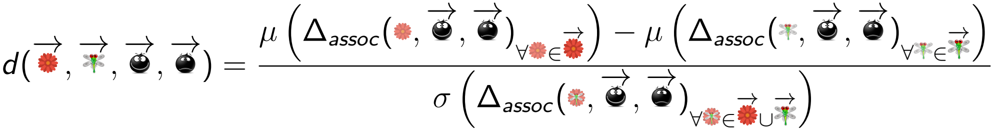
Languages

Arabic (ar), Catalan (ca), Croatian (hr), English (en), German (de), Italian (it), Russian (ru), Spanish (es) and Turkish (tr)

Experiments

What we Report here (More in the Paper!)

- Size effect



Sawilowsky's scale: very small (d<0.01), small (<0.20), medium (<0.50), large (<0.80), very large (<1.20), and huge (<2.00)

Experiments

What we Report here (More in the Paper!)

- Size effect

$$d(\vec{\text{🌺}}, \vec{\text{🦟}}, \vec{\text{😀}}, \vec{\text{☹}}) = \frac{\mu\left(\Delta_{assoc}(\text{🌺}, \vec{\text{😀}}, \vec{\text{☹}})_{\forall \text{🌺} \in \vec{\text{🌺}}}\right) - \mu\left(\Delta_{assoc}(\text{🦟}, \vec{\text{😀}}, \vec{\text{☹}})_{\forall \text{🦟} \in \vec{\text{🦟}}}\right)}{\sigma\left(\Delta_{assoc}(\text{🌺}, \vec{\text{😀}}, \vec{\text{☹}})_{\forall \text{🌺} \in \vec{\text{🌺}} \cup \vec{\text{🦟}}}\right)}$$

Sawilowsky's scale: very small (d<0.01), small (<0.20), medium (<0.50), large (<0.80), very large (<1.20), and huge (<2.00)

  - CA-WEAT: median and 95% CI with order statistics
  - WEAT, CA-WEAT, X-WEAT: 5,000 bootstraps (median and 95% CI)

- Size effect

$$d(\vec{\text{🌼}}, \vec{\text{🦟}}, \vec{\text{😀}}, \vec{\text{☹}}) = \frac{\mu\left(\Delta_{assoc}(\text{🌼}, \vec{\text{😀}}, \vec{\text{☹}})_{\forall \text{🌼} \in \vec{\text{🌼}}}\right) - \mu\left(\Delta_{assoc}(\text{🦟}, \vec{\text{😀}}, \vec{\text{☹}})_{\forall \text{🦟} \in \vec{\text{🦟}}}\right)}{\sigma\left(\Delta_{assoc}(\text{🌼🦟}, \vec{\text{😀}}, \vec{\text{☹}})_{\forall \text{🌼🦟} \in \vec{\text{🌼}} \cup \vec{\text{🦟}}}\right)}$$

Sawilowsky's scale: very small (d<0.01), small (<0.20), medium (<0.50), large (<0.80), very large (<1.20), and huge (<2.00)

  - CA-WEAT: median and 95% CI with order statistics
  - WEAT, CA-WEAT, X-WEAT: 5,000 bootstraps (median and 95% CI)

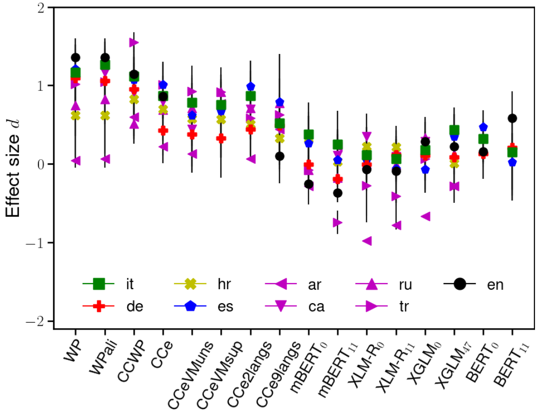
- IAT1 (🌼 🦟); IAT2 (🎸 🔫) is equivalent

Experiments

Do our embeddings show (human) biases?
All embedding models? All languages?

Experiments

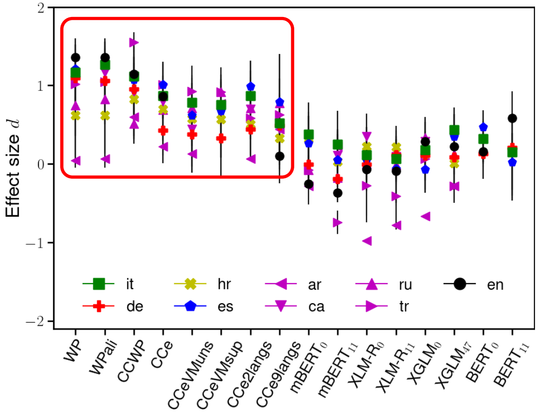
Wide Overview (WEAT, CA-WEAT)



Wide Overview (WEAT, CA-WEAT)

Word embeddings:

- All WE models have $d > 0$

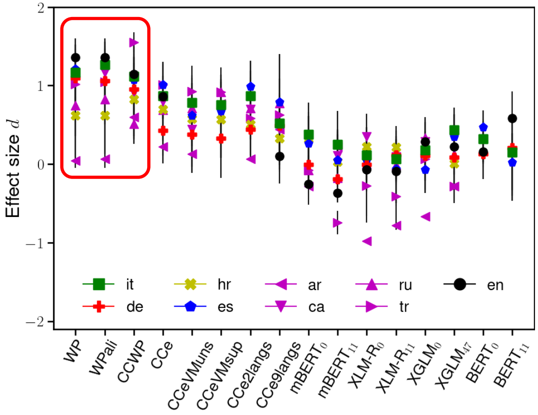


Experiments

Wide Overview (WEAT, CA-WEAT)

Word embeddings:

- All WE models have $d > 0$
- Pre-trained models have higher σ across languages

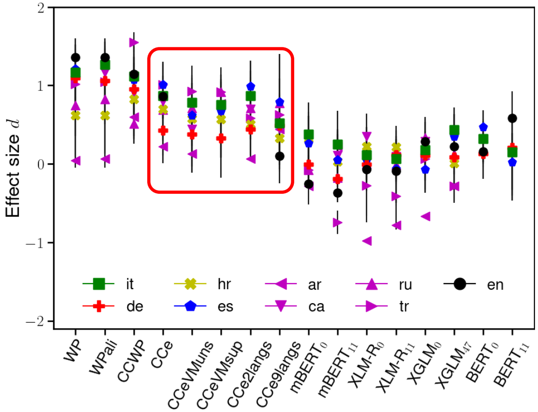


Experiments

Wide Overview (WEAT, CA-WEAT)

Word embeddings:

- All WE models have $d > 0$
- Pre-trained models have higher σ across languages
- Equivalent projection methods

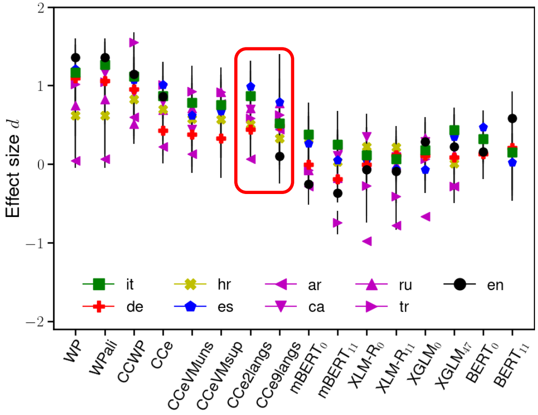


Experiments

Wide Overview (WEAT, CA-WEAT)

Word embeddings:

- All WE models have $d > 0$
- Pre-trained models have higher σ across languages
- Equivalent projection methods
- Multilinguality attenuates

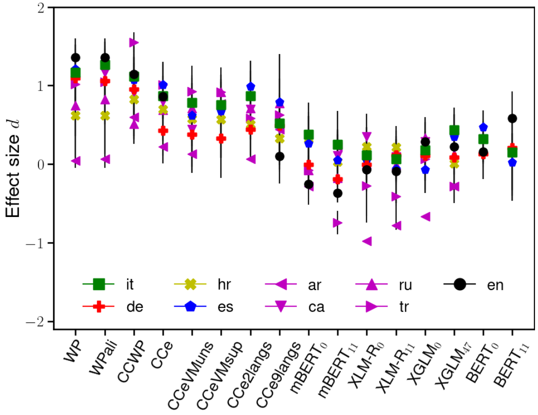


Experiments

Wide Overview (WEAT, CA-WEAT)

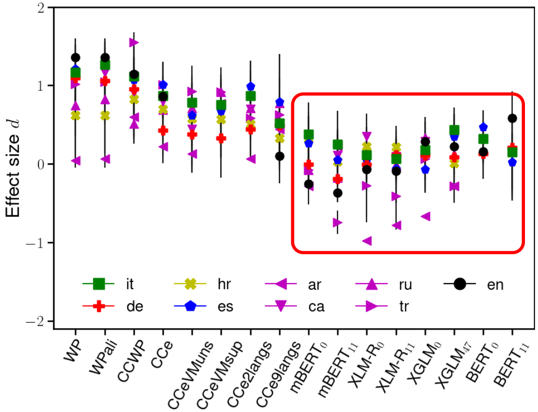
Word embeddings:

- All WE models have $d > 0$
- Pre-trained models have higher σ across languages
- Equivalent projection methods
- Multilinguality attenuates
- No universal d



Contextual embeddings:

- d compatible with no bias

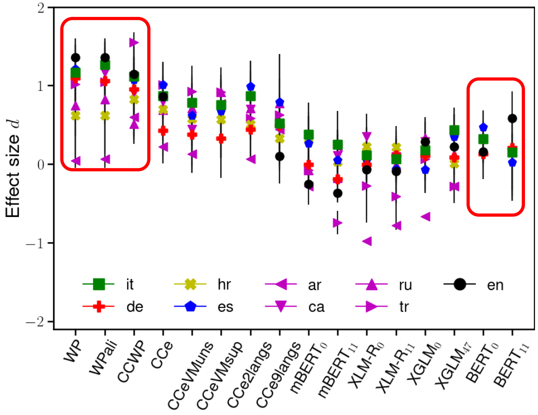


Experiments

Wide Overview (WEAT, CA-WEAT)

Contextual embeddings:

- *d* compatible with no bias
- Effect of *contextualisation*

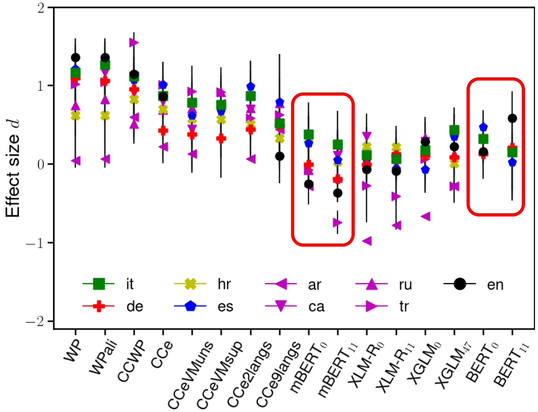


Experiments

Wide Overview (WEAT, CA-WEAT)

Contextual embeddings:

- d compatible with no bias
- Effect of *contextualisation*
- But multilinguality attenuates further

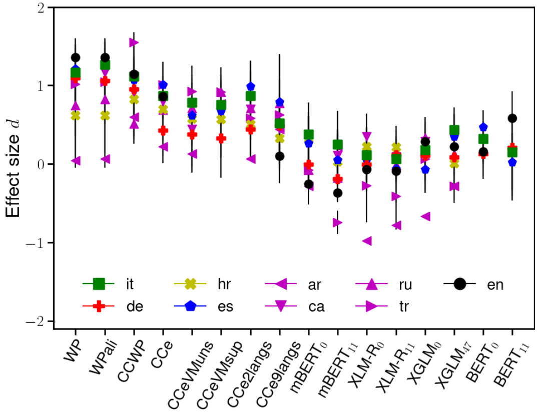


Experiments

Wide Overview (WEAT, CA-WEAT)

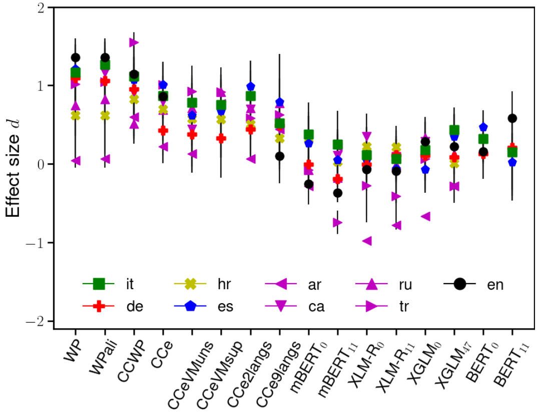
Contextual embeddings:

- d compatible with no bias
- Effect of *contextualisation*
- But multilinguality attenuates further
- No universal d



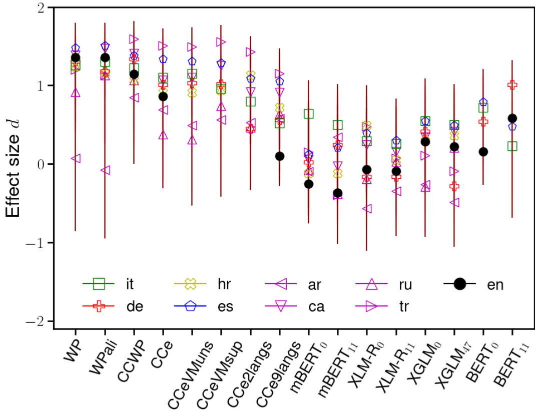
Experiments

Wide Overview (CA-WEAT vs X-WEAT)



Experiments

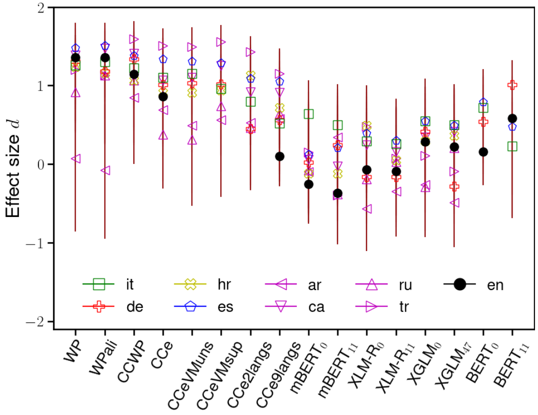
Wide Overview (CA-WEAT vs X-WEAT)



Experiments

Wide Overview (CA-WEAT vs X-WEAT)

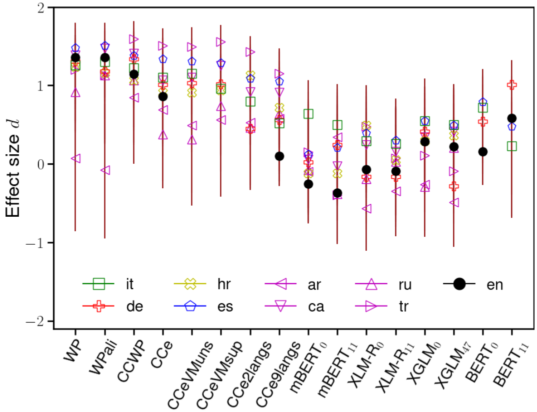
- X-WEAT shows similar trends as CA-WEAT
- But! With a higher dispersion across languages



Experiments

Wide Overview (CA-WEAT vs X-WEAT)

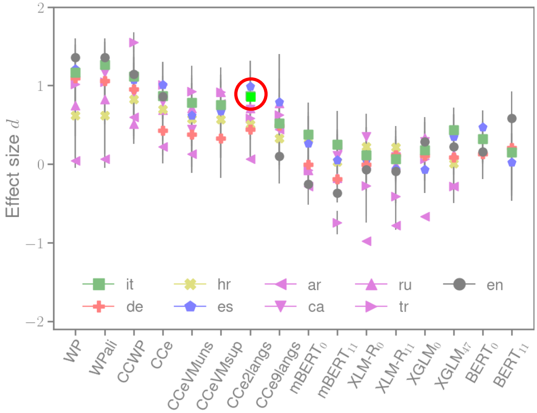
- X-WEAT shows similar trends as CA-WEAT
- But! With a higher dispersion across languages
- No universal *d*



Experiments

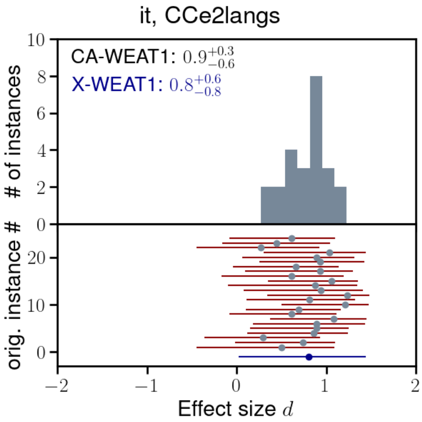
Wide Overview (CA-WEAT vs X-WEAT)

Let's focus!



Experiments

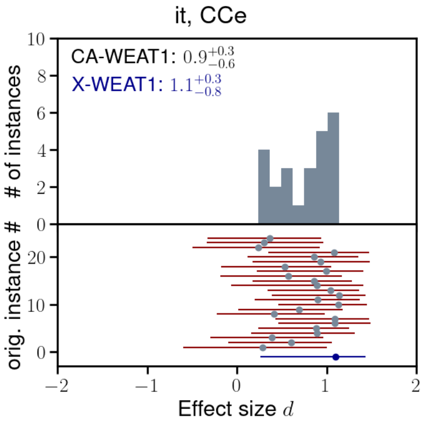
WEAT vs X-WEAT vs CA-WEAT



- Lists show a high dispersion (bootstrapped and averaged)
- X-WEAT lies within CA-WEAT (close cultures?)

Experiments

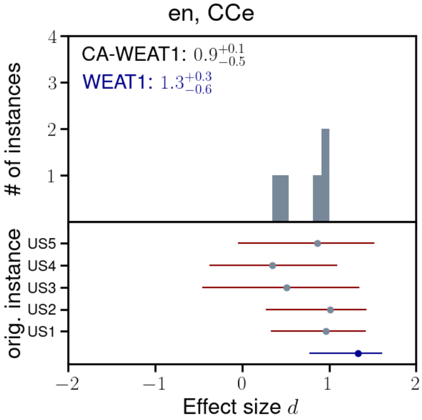
WEAT vs X-WEAT vs CA-WEAT



- Lists show a high dispersion (bootstrapped and averaged)
- X-WEAT lies within CA-WEAT (close cultures?)
- Distributions non-normal (yet!)

Experiments

WEAT vs X-WEAT vs CA-WEAT



- Lists show a high dispersion (bootstrapped and averaged)
- X-WEAT lies within CA-WEAT (close cultures?)
- Distributions non-normal (yet!)
- English interesting for further study (and Spanish, and French... :-))

Experiments

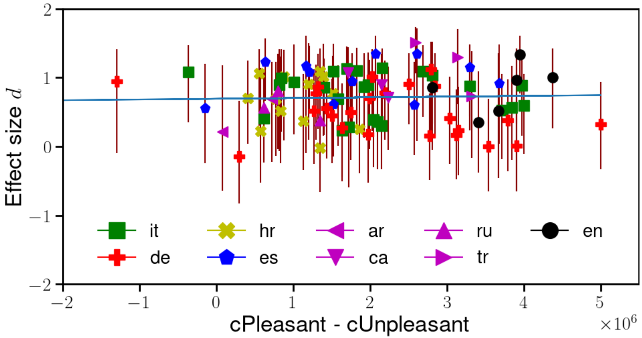
Why is d non-universal?

Is it data differences? Is it forcing multiliguality?
Is it the dispersion?

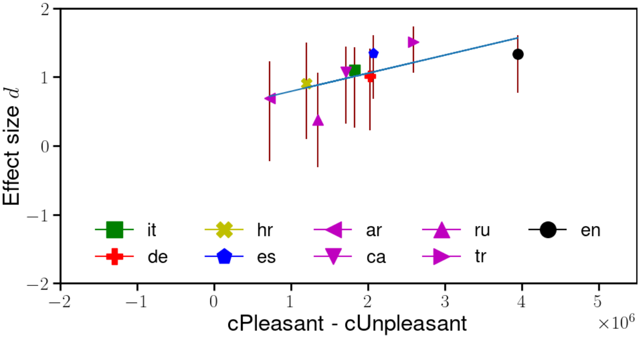
Experiments

Asymmetries in Concepts Frequencies (CCe)

WEAT1+X-WEAT1+CA-WEAT1: no relation

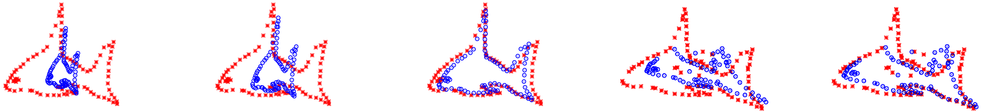


X-WEAT1: Simpson's paradox?



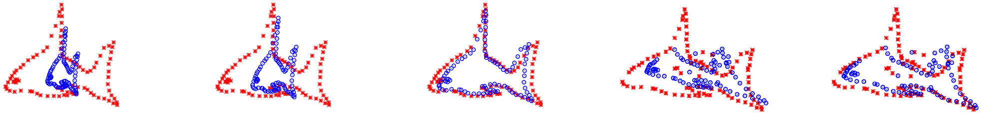
Experiments

Isomorphism



Experiments

Isomorphism



- Measures: Gromov-Hausdorff distance and Eigenvector similarity
- Isomorphism between a language (sub-)space and the English (sub-)space
- For contextual models we consider the vocab from CCe

Experiments

Isomorphism between a Language (sub-)Space and the English (sub-)Space

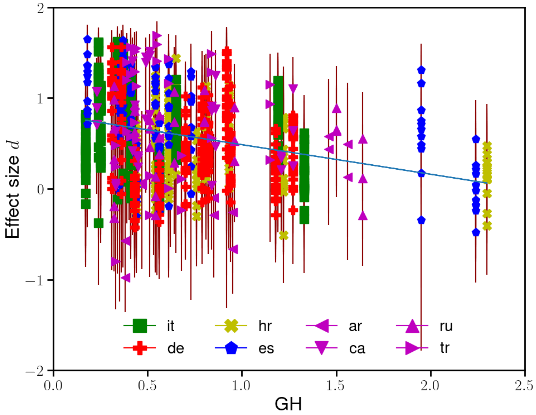
	ar		ca		de		es		hr		it		ru		tr	
	EV	GH	EV	GH	EV	GH	EV	GH	EV	GH	EV	GH	EV	GH	EV	GH
WP	106	0.47	12	0.49	12	0.31	10	0.18	42	0.54	21	0.24	16	0.43	49	0.39
WPali	143	0.55	22	0.51	22	0.36	16	0.37	46	0.61	19	0.34	30	0.32	36	0.44
CCWP	15	0.40	85	0.42	42	0.92	23	0.41	51	0.65	41	0.37	32	0.64	28	0.55
CCe	55	0.62	253	0.23	26	0.79	166	0.54	91	0.61	223	0.25	8	0.56	25	0.43
CCeVMuns	229	1.56	229	1.27	27	0.82	167	1.95	69	0.93	220	1.19	27	0.96	36	0.84
CCeVMsup	36	0.56	231	0.86	32	0.70	87	0.73	27	0.61	123	0.65	25	0.80	11	0.41
CCe2langs	93	0.53	8	0.43	19	0.94	72	0.35	33	0.81	51	0.41	39	0.51	64	0.61
CCe9langs	475	1.46	23	0.84	171	1.27	21	0.61	53	1.22	51	0.41	403	1.50	149	1.15
mBERT_0	154	0.85	133	0.33	95	0.56	99	0.56	270	0.44	131	0.17	161	0.54	589	0.51
XLM-R_0	54	0.38	74	0.45	59	0.43	150	0.44	58	0.54	113	0.56	111	0.32	277	0.33
XGLM_0	67	0.95	88	1.21	144	1.18	135	2.24	*2584	*2.30	130	1.33	85	1.64	475	0.68

- No clear distinction between WE and CE wrt. isomorphism distances
- Language and embedding model effects are also mixed

Experiments

Isomorphism between a Language (sub-)Space and the English (sub-)Space

- correlation(GH, d)=-0.29; describes a 10% of the variance



Outline

1 What is a Bias and how do we Measure them
- IAT: Implicit Association Tests
- WEAT: Association Tests in Word Embeddings

2 Multilinguality and Cultural-Aware WEAT (CA-WEAT)

3 Experiments
- Wide Overview
- WEAT vs X-WEAT vs CA-WEAT
- Data Asymmetries and Isomorphism

4 Conclusions

Conclusions

Wrapping up

- Using (literal) translation in NLP does not in general preserve culture
- We therefore create CA-WEAT (in contrast to X-WEAT) to analyse desirable biases in embeddings across languages

Conclusions

Wrapping up

- Using (literal) translation in NLP does not in general preserve culture
- We therefore create CA-WEAT (in contrast to X-WEAT) to analyse desirable biases in embeddings across languages

- Monolingual and bilingual WE reproduce non-social human biases
- We do not observe a universal value even in the comparable setting
- Contextualisation and multiliguality attenuate biases, why?

Conclusions

Wrapping up

- Using (literal) translation in NLP does not in general preserve culture
- We therefore create CA-WEAT (in contrast to X-WEAT) to analyse desirable biases in embeddings across languages

- Monolingual and bilingual WE reproduce non-social human biases
- We do not observe a universal value even in the comparable setting
- Contextualisation and multiliguality attenuate biases, why?

- Due to the large variablility (models & languages) we want...

Conclusions

Future Work

- Better understanding of individual vs cultural differences
- Better understanding of intralanguage cultural differences
- Better understanding of language models

Future Work

- Better understanding of individual vs cultural differences
- Better understanding of intralanguage cultural differences
- Better understanding of language models

...time for a short digression?

Conclusions

A Reviewer's Comment

There is a huge variability.

Shouldn't one use more (WEAT) tests?

How do we find more tests?!

We want universality...

Conclusions

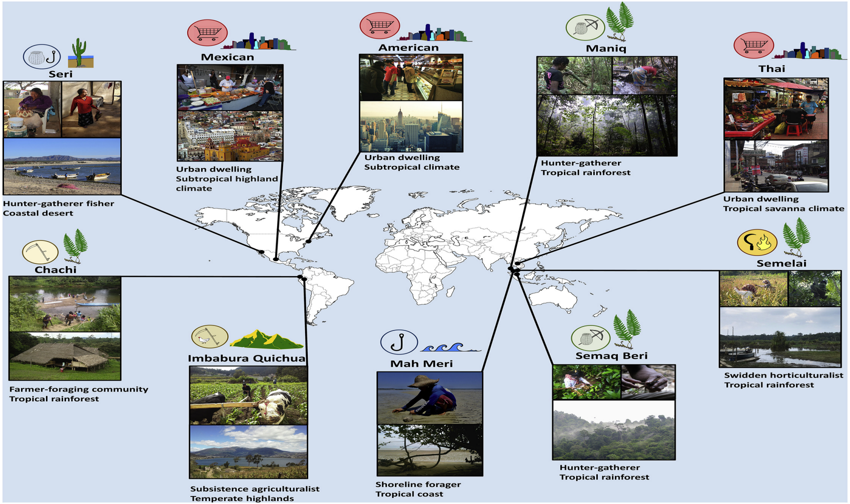
The Perception of Odor Pleasantness is Shared Across Cultures

[Arshamian et al., Current Biology, 2022]

- Culture plays a minimal role in the perception of odor pleasantness
- Individuals within cultures vary as to which odors they find pleasant
- Human olfactory perception is strongly constrained by universal principles

Conclusions

The Perception of Odor Pleasantness is Shared Across Cultures



Conclusions

Back into the Future Work!

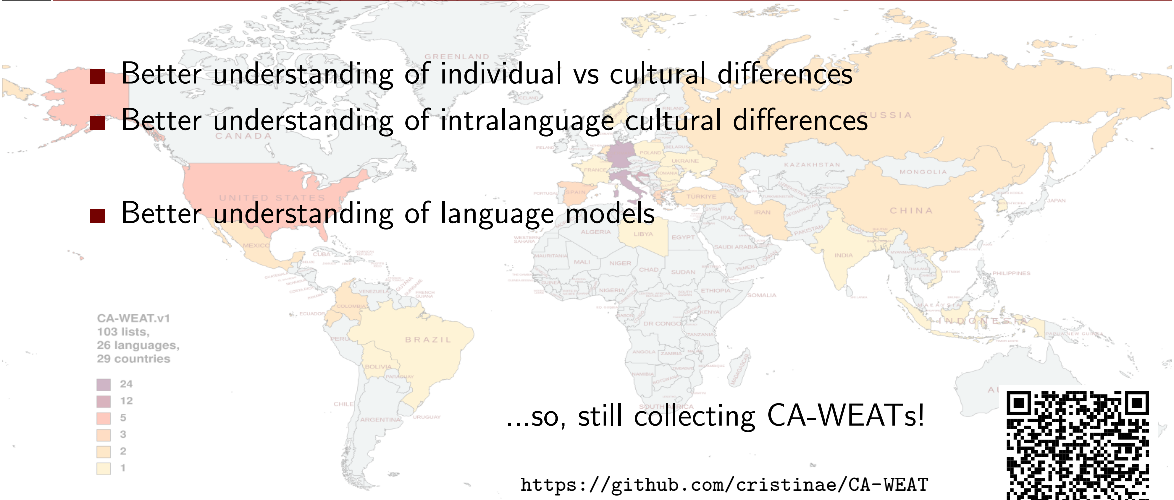
- Better understanding of individual vs cultural differences
- Better understanding of intralanguage cultural differences
- Better understanding of language models

Conclusions

Future Work

- Better understanding of individual vs cultural differences
- Better understanding of intralanguage cultural differences

- Better understanding of language models



...so, still collecting CA-WEATs!

`https://github.com/cristinae/CA-WEAT`

That’s All Folks!

Thanks! And...

That’s All Folks!

Datasets Related to Multilinguality and Cultural Diversity

ArtELingo: A Million Emotion Annotations of WikiArt with Emphasis on Diversity over Language and Culture (Mohamed et al., EMNLP 2022)

Examples from ArtELingo

	شلال طبيعي جميل. مشاعر النمو والحيوية والطاقة موجودة. Translation: Beautiful natural waterfall. Feelings of growth, vitality and energy. **Excitement** **Arabic**	The water that's rushing downward looks like a bride's wedding veil. **Awe** **English**	瀑布就像四蹄生风的白马如潮水涌来，非常的壮观 Translation: The waterfall is like a white horse and wind, it is spectacular. **Contentment** **Chinese**

Crossmodal-3600: A Massively Multilingual Multimodal Evaluation Dataset (Thapliyal et al., EMNLP 2022)

Source: *Porsche Museum, Stuttgart* by *Brian Solis*.

English	• A vintage sports car in a showroom with many other vintage sports cars • The branded classic cars in a row at display
Spanish	• Automóvil clásico deportivo en exhibición de automóviles de galería *(Classic sports car in gallery car display)* • Coche pequeño de carreras color plateado con el número 42 en una exhibición de coches *(Small silver racing car with the number 42 at a car show)*
Thai	• รถเปิดประทุนหลายสีจอดเรียงกันในที่จัดแสดง *(Multicolored convertibles line up in the exhibit)* • รถแข่งวินเทจจอดเรียงกันหลายคันในงานจัดแสดง *(Several vintage racing cars line up at the show)*

That's All Folks!

Datasets Related to Multilinguality and Cultural Diversity

Figure 2: A sample of images in the XM3600 dataset, together with the language for which they have been selected. Overall, the images span regions over 36 different languages and 6 different continents.

That’s All Folks!

Datasets Related to Multilinguality and Cultural Diversity

Stanceosaurus: Classifying Stance Towards Multicultural Misinformation (Zheng et al., EMNLP 2022)

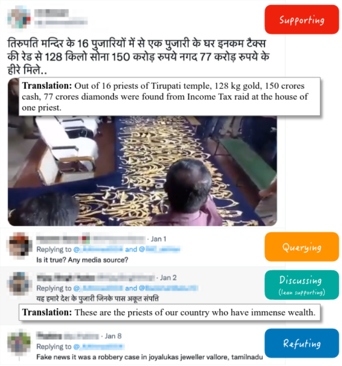


Figure 1: Example Hindi and English tweets in Stanceosaurus with stance towards the claim *“Raid at Tirupati temple priest’s house, 128 kg gold found”*.