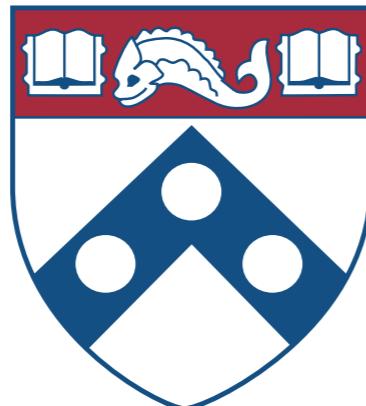


The Promise of Crowdsourcing for Machine Learning and Data Science

Chris Callison-Burch

August 4, 2020



Penn
UNIVERSITY *of* PENNSYLVANIA

Amazon Mechanical Turk

Access a global, on-demand, 24x7 workforce

Amazon Mechanical Turk (MTurk) is a crowdsourcing marketplace that makes it easier for individuals and businesses to outsource their processes and jobs to a distributed workforce who can perform these tasks virtually. This could include anything from conducting simple data validation and research to more subjective tasks like survey participation, content moderation, and more. MTurk enables companies to harness the collective intelligence, skills, and insights from a global workforce to streamline business processes, augment data collection and analysis, and accelerate machine learning development.

While technology continues to improve, there are still many things that human beings can do much more effectively than computers, such as moderating content, performing data deduplication, or research. Traditionally, tasks like this have been accomplished by hiring a large temporary workforce, which is time consuming, expensive and difficult to scale, or have gone undone. Crowdsourcing is a good way to break down a manual, time-consuming project into smaller, more manageable tasks to be completed by distributed workers over the Internet (also known as ‘microtasks’).

Optimize efficiency

MTurk is well-suited to take on simple and repetitive tasks in your workflows which need to be handled manually. Using MTurk to outsource microtasks ensures that work gets done quickly, while freeing up time and resources for the company – so internal staff can focus on higher value activities.

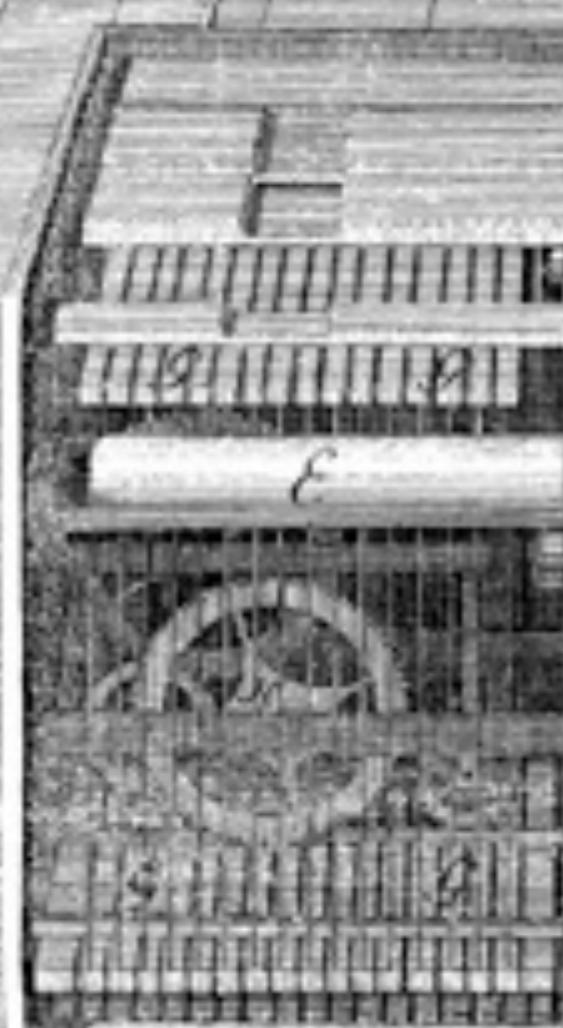
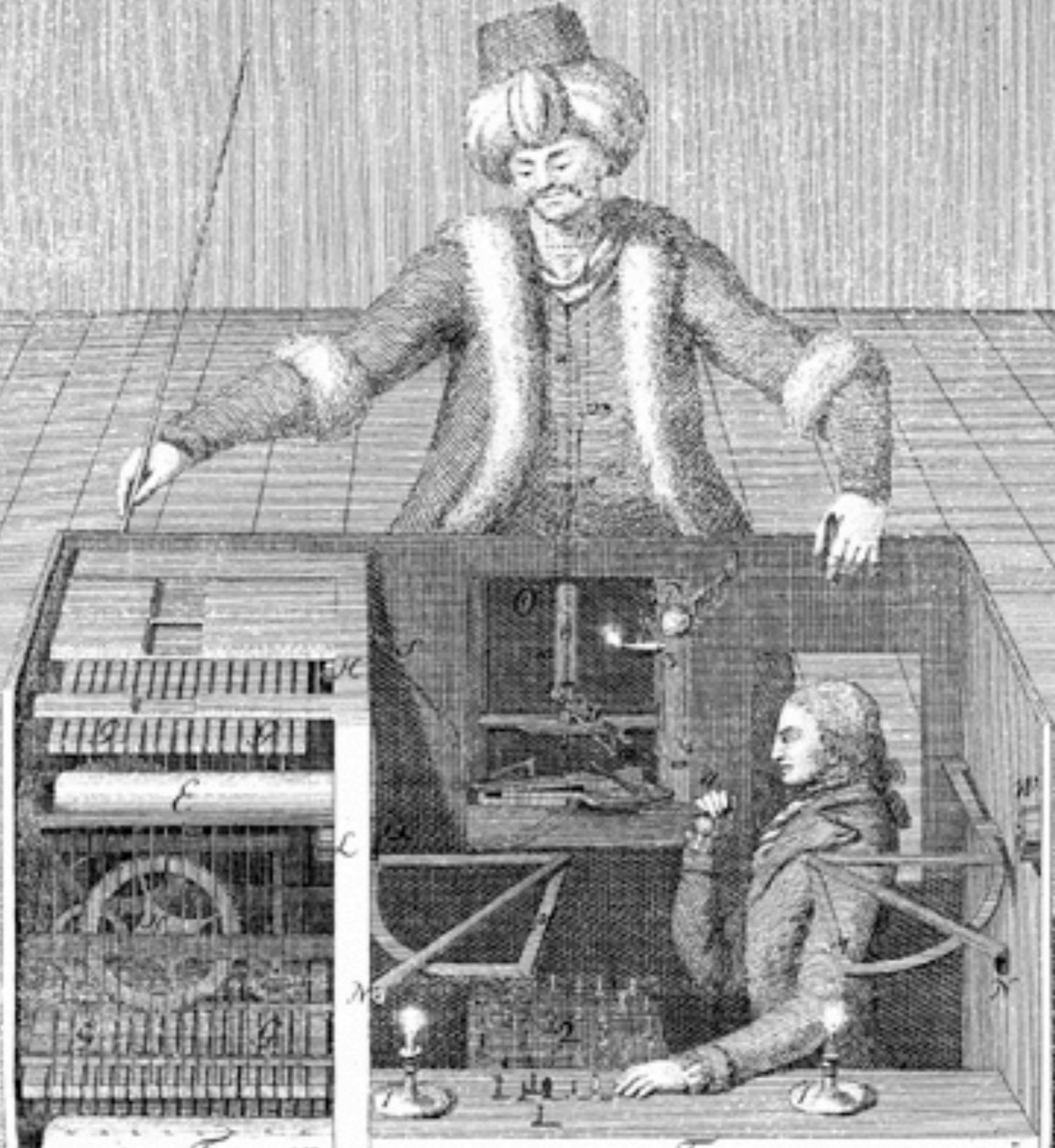
Increase flexibility

Scaling up and down a workforce isn’t the easiest undertaking. With access to a global, on-demand, 24x7 workforce, MTurk enables businesses and organizations to get work done easily and quickly when they need it – without the difficulty associated with dynamically scaling your in-house workforce.

Reduce cost

MTurk offers a way to effectively manage labor and overhead costs associated with hiring and managing a temporary workforce. By leveraging the skills of distributed Workers on a pay-per-task model, you can significantly lower costs while achieving results that might not have been possible with just a dedicated team.





[All HITs](#) | [HITs Available To You](#) | [HITs Assigned To You](#)Search for **HITs** containing that pay at least \$ 0.00 for which you are qua**Timer:** 00:00:00 of 5 minutes Want to work on this HIT? Want to see other HITs?**Total Earned:** Unavaila**Accept HIT****Skip HIT****Total HITs Submitted:** 0

Enter Postmark & Stamp Information for a Postcard

Requester: Cardcow**Reward:** \$0.01 per HIT**HITs Available:** 2**Duration:** 5 minutes**Qualifications Required:** Data Entry for Postcards has been granted**Enter Postmark & Stamp Information for this card****Postmark City:****Postmark State:
(or Country)****Postmark Date:
(Ex: Nov-09)****Postmark Year:
(Ex: 1909)****Stamp: (Ex: 1c, 2c,
half penny)**

(month & day)

[All HITs](#) | [HITs Available To You](#) | [HITs Assigned To You](#)Search for HITs containing that pay at least \$ 0.00 for which you are qu:Timer: 00:00:00 of 5 minutes Want to work on this HIT? [Accept HIT](#) Want to see other HITs? [Skip HIT](#)

Total Earned: Unavaila

Total HITs Submitted: 0

Enter Postmark & Stamp Information for a Postcard

Requester: Cardcow

Reward: \$0.01 per HIT

HITs Available: 2

Duration: 5 minutes

A reward has been granted

**Postmark City:**

Barre

Postmark State:

MA

Postmark Date:

Oct-11

Postmark Year:

1886

Stamp:

1c

Postmark City:

Postmark State:
(or Country)Postmark Date:
(Ex: Nov-09)

\$ 0.01

(month & day)



In 2006 Fei-Fei Li was a new CS professor at UIUC. Everyone was trying to develop better algorithms that would make better decisions, regardless of the data.



But she realized a limitation to this approach—the best algorithm wouldn't work well if the data it learned from didn't reflect the real world.
Her solution: build a better dataset.



“We decided we wanted to do something that was completely historically unprecedented. We’re going to map out the entire world of objects.” The resulting dataset was called ImageNet.

IMAGENET



MTurk for Natural Language Processing

Snow, O'Connor, Jurafsky and Ng's EMNLP 2008 paper pioneered the use of Mechanical Turk for NLP

Affect Recognition

fear("Tropical storm forms in Atlantic") >
fear("Goal delight for Beckham")

Word Similarity

sim(lad, boy) > sim(rooster, noon)

Textual Entailment

if "Microsoft was established in Italy in 1985"
then "Microsoft was established in 1985"?

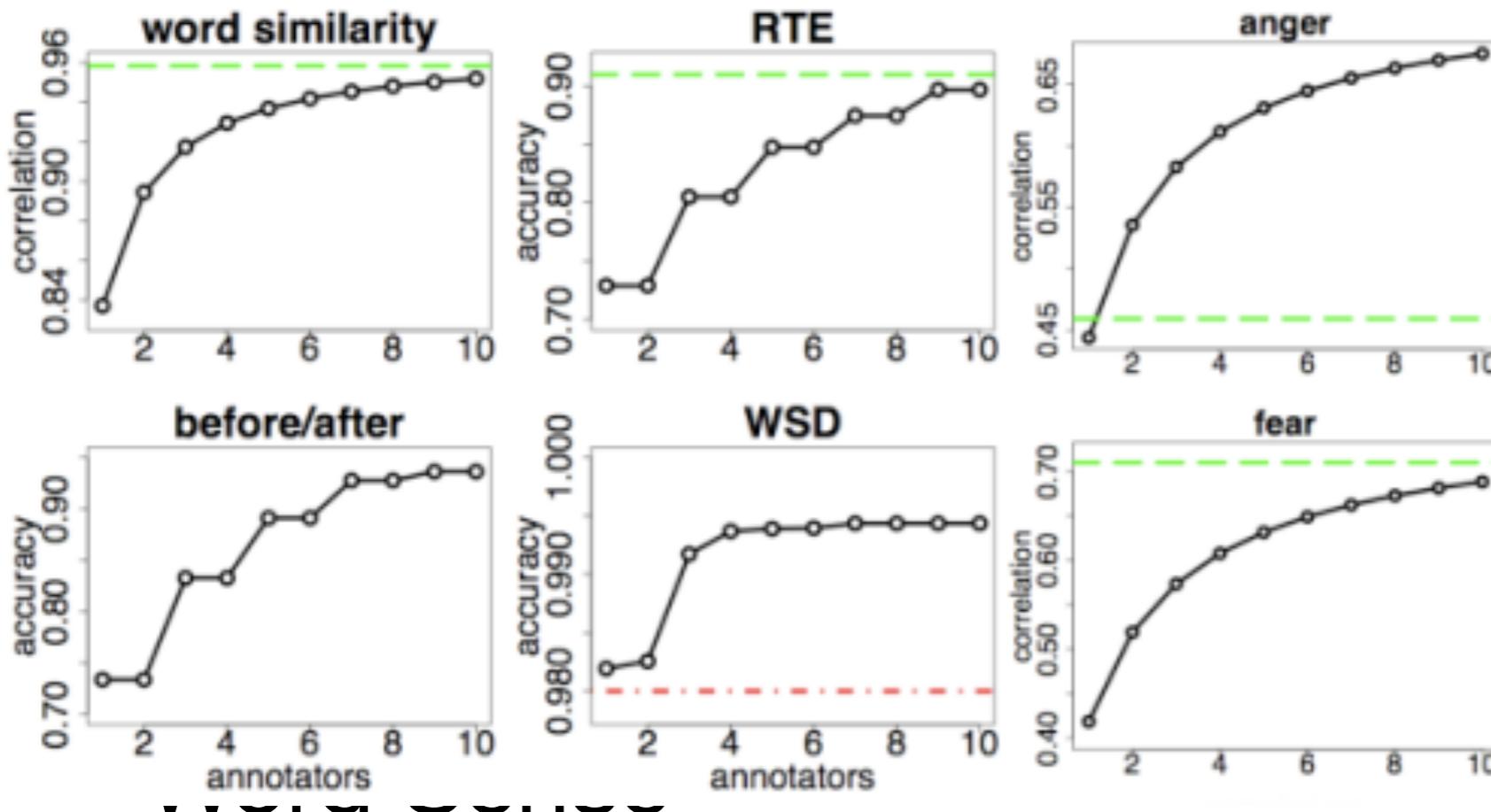
Word Sense

"a bass on the line" v. "a funky bass line"

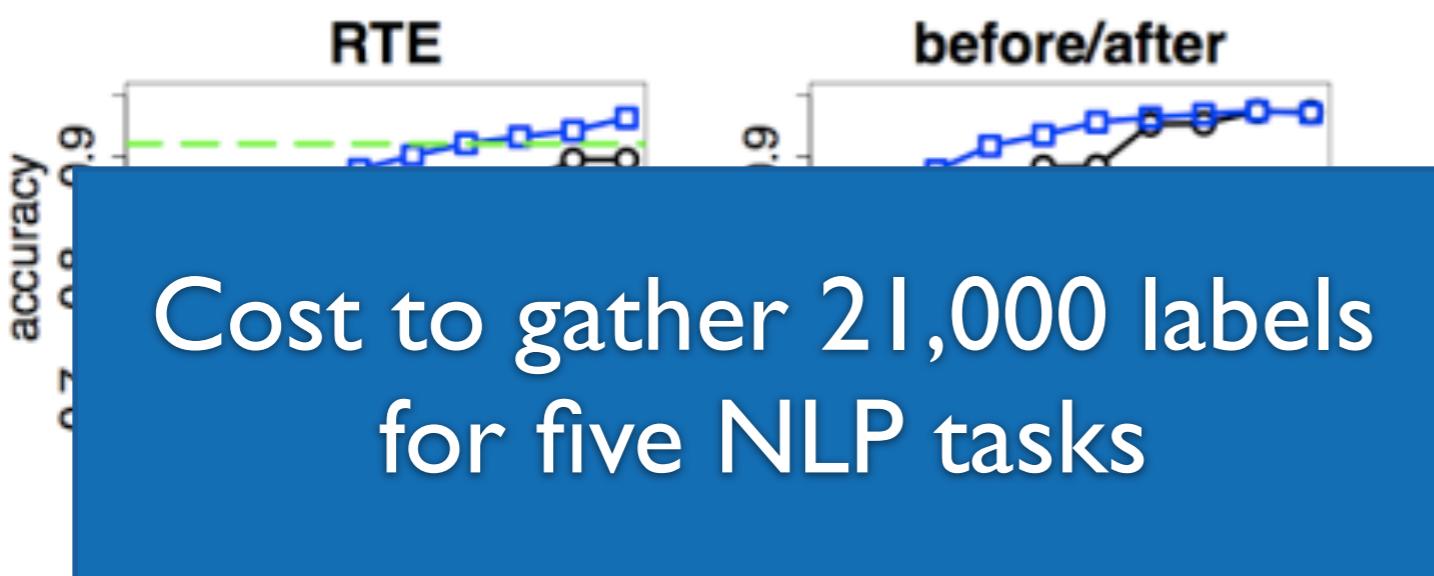
Temporal Annotation

ran happens before *fell* in:
"The horse raced past the barn fell."

NLP Annotation



Combine non-expert judgments for high correlation with experts



Weight w/ small amount of gold standard data for

\$ 25.82

\$100 Challenge



- Workshop on Using Mechanical Turk for Speech and Language Applications
- 35 researchers spent \$100, wrote papers, and distributed their data
- Mark Dredze and I wrote an overview paper digesting the results

Who are the Turkers?

- Requesters are given very little information about Turkers - basically just a serial number
- No names, no demographic information (like what languages they speak)
- Cannot assume that they have a particular set of skills
- They should be treated as non-experts
- Quality control is a major challenge
- It important to design tasks to be simple and easy to understand
- Why would Turkers work for pennies?

Machine Translation

- Translation rules are learned from bilingual texts using machine learning techniques

Arabic

فالتعذيب لا يزال يمارس على نطاق واسع

وتتم عمليات الاعتقال والاحتجاز دون سبب بصورة
روتينية

وحان وقت التحلّى بال بصيرة والشجاعة السياسية .

...

English

Torture is still being practised on a wide scale.

Arrest and detention without cause take place routinely.

This is a time for vision and political courage

...

Chinese

我国 能源 原材料 工业 生产 大幅度 增长 .

非国大 要求 阻止 更多 被 拘留 人员 死亡 .

...

English

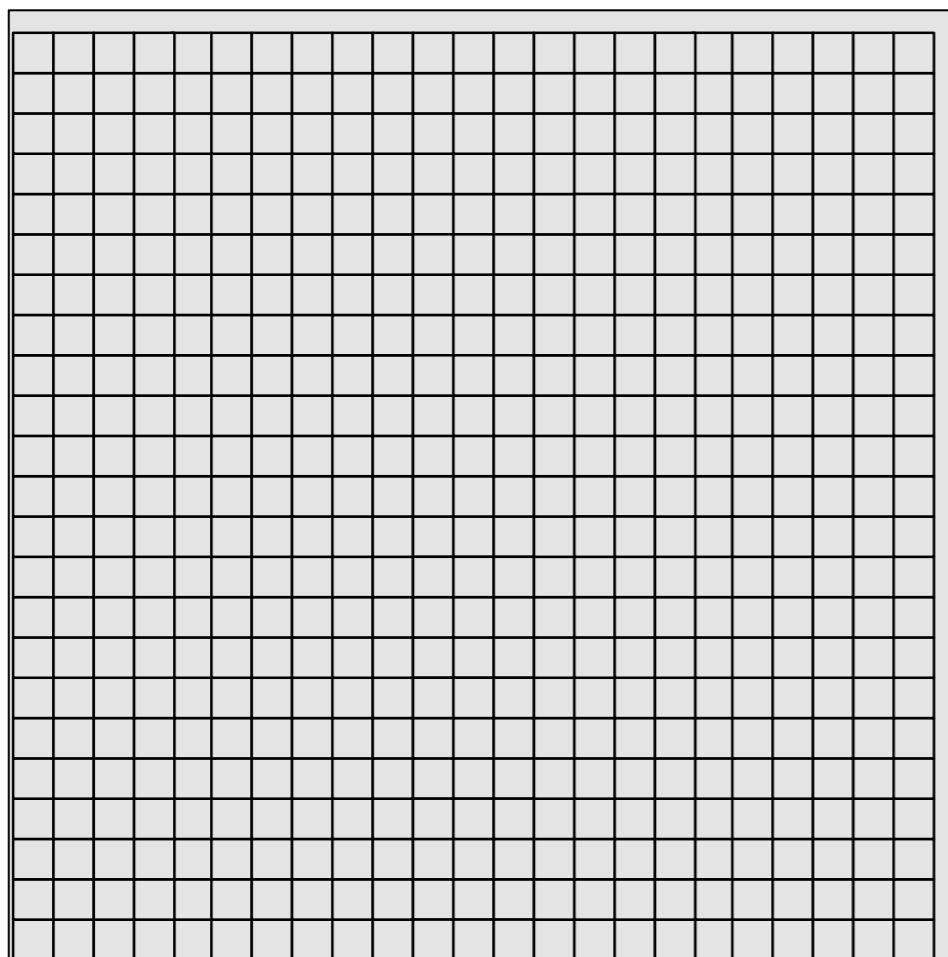
China's energy and raw materials production up.

ANC calls for steps to prevent deaths in police custody .

...

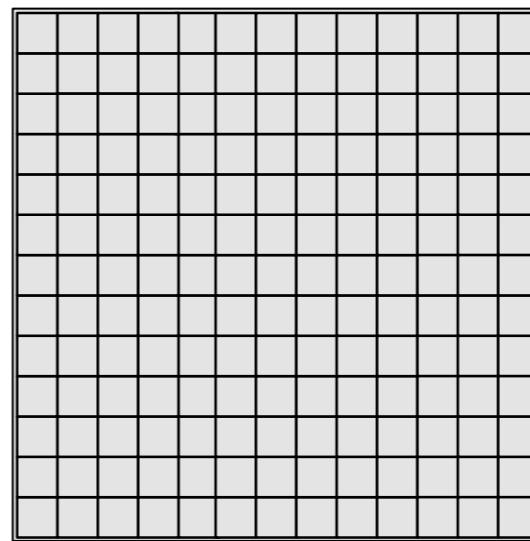
Training data varies by language

1000M



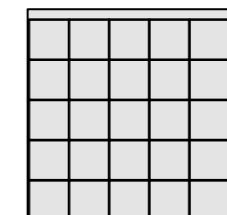
French-English
 10^9 word webcrawl

200M



Arabic and Chinese
DARPA GALE

50M

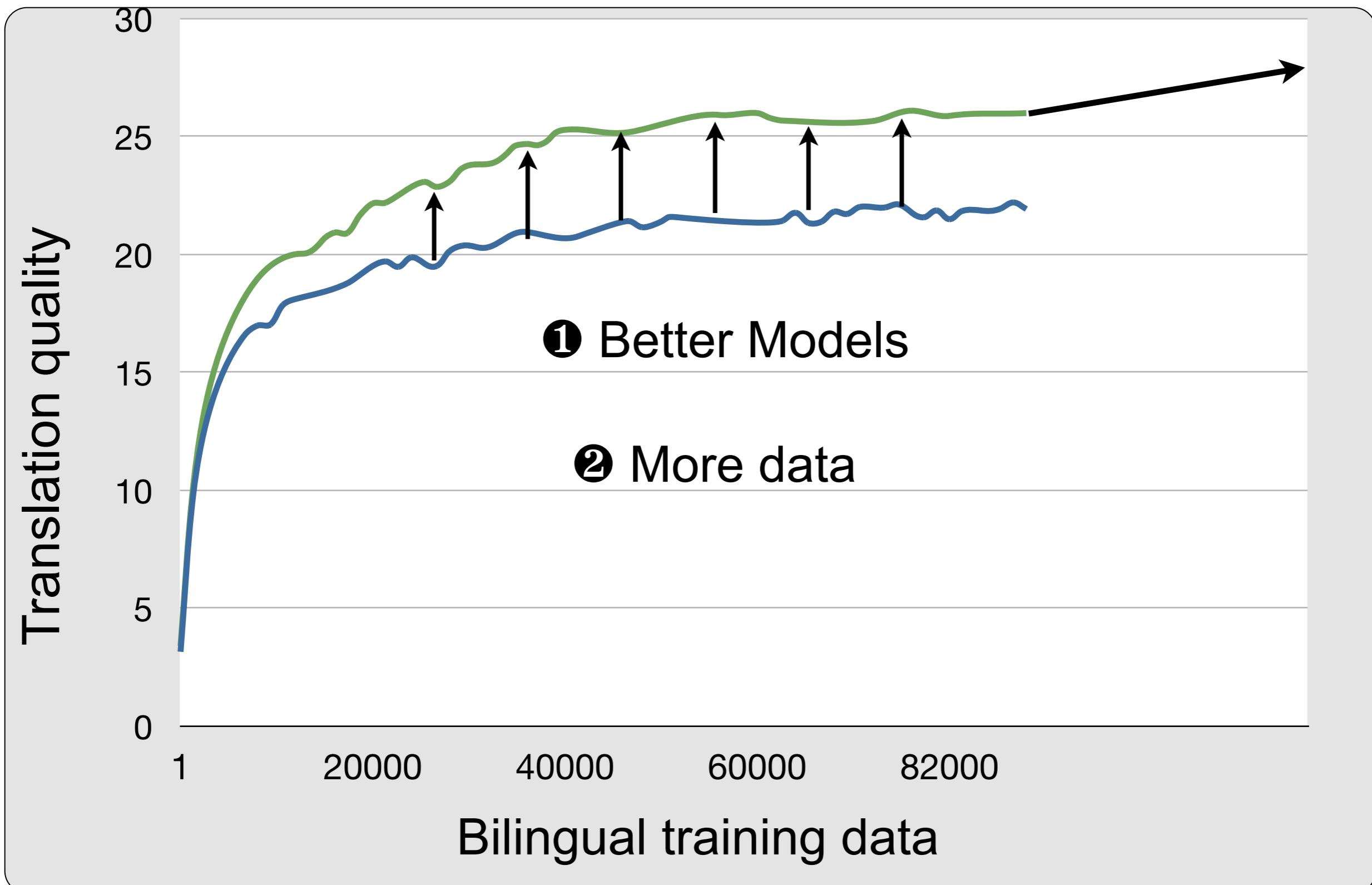


European
Parliament

1.5M

Urdu

How to Improve Machine



Can Turkers Translate?

Urdu

میں اس خطے میں ابتدائی 1994 انسانوں کی باقیات جو تقریباً 8 لاکھ سال پرانی مانی جاتی ہے، دریافت کی گئیں جنہیں ہومو اینٹی سیسٹر یعنی 'بانی انسان' کا نام دیا گیا۔

اس سے قبل 6 لاکھ پرانے انسان جنہیں سائنسی اصطلاح میں ہومو ہیڈلبرجنیس کہا جاتا ہے، اس خطے کے قدیم ترین رہائشی مانے جاتے تھے۔

آثارِ قدیمه کے ماہرین کا کہنا ہے کہ انہیں ایسے شواہد ملے ہیں جن سے پتہ چلتا ہے کہ اس خطے کے لوگ ڈھلانی کیسے ہوئے اور اس بھی

LDC Translation

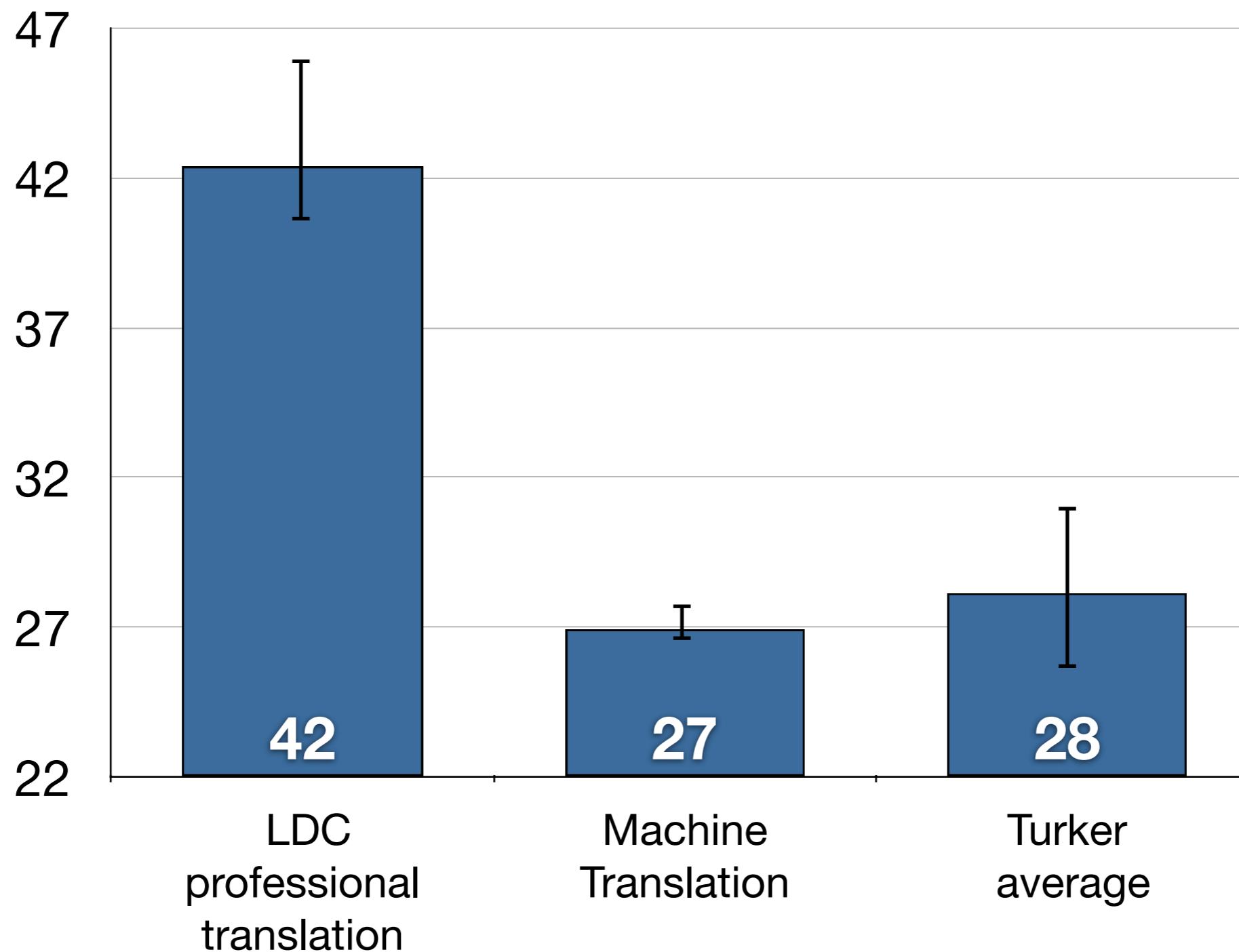
In 1994, the remains of early human beings who were believed to be eight hundred thousand years old were discovered who were given the name homo antecessor meaning the 'founder man'. Prior to this, the six hundred thousand years old man, called homo heidelbergensis in scientific terms, was believed to be the earliest resident of this area.

Turk Translation

In 1994, the remains of pre-historic man, which are believed to be 800,000 years old were discovered and they were named 'Home Antecessor' meaning 'The Founding Man'. Prior to that 6 lac years old humans, named as Homogenisens in scientific terms, were believed to be the oldest dwellers of this area.

Professionals versus Turkers*

*Without Quality Control



Avoiding dieting to prevent from flu	abstention from dieting in order to avoid Flu	Abstain from decrease eating in order to escape from flue	In order to be safer from flu quit dieting
This research of American scientists came in front after experimenting on mice.	This research from the American Scientists have come up after the experiments on rats.	This research of American scientists was shown after many experiments on mouses.	According to the American Scientist this research has come out after much experimentations on rats.
Experiments proved that mice on a lower calorie diet had comparatively less ability to fight the flu virus.	It has been proven from experiments that rats put on diet with less calories had less ability to resist the Flu virus.	It was proved by experiments the low calories eaters mouses had low defending power for flue in ratio.	Experimentaions have proved that those rats on less calories diet have developed a tendency of not overcoming the flu virus.
research has proven this old myth wrong that its better to fast during fever.	Research disproved the old axiom that " It is better to fast during fever"	The research proved this old talk that decrease eating is useful in fever.	This Research has proved the very old saying wrong that it is good to starve while in fever.

Avoiding dieting to prevent from flu	abstention from dieting in order to avoid Flu	Abstain from decrease eating in order to escape from flue	In order to be safer from flu quit dieting
This research of American scientists came in front after experimenting on mice.	This research from the American Scientists have come up after the experiments on rats.	This research of American scientists was shown after many experiments on mouses.	According to the American Scientist this research has come out after much experimentations on rats.
Experiments proved that mice on a lower calorie diet had comparatively less ability to fight the flu virus.	It has been proven from experiments that rats put on diet with less calories had less ability to resist the Flu virus.	It was proved by experiments the low calories eaters mouses had low defending power for flue in ratio.	Experimentaions have proved that those rats on less calories diet have developed a tendency of not overcoming the flu virus.
research has proven this old myth wrong that its better to fast during fever.	Research disproved the old axiom that " It is better to fast during fever"	The research proved this old talk that decrease eating is useful in fever.	This Research has proved the very old saying wrong that it is good to starve while in fever.

Avoiding dieting to prevent from flu	abstention from dieting in order to avoid Flu	Abstain from decrease eating in order to escape from flue	In order to be safer from flu quit dieting
This research of American scientists came in front after experimenting on mice.	This research from the American Scientists have come up after the experiments on rats.	This research of American scientists was shown after many experiments on mouses.	According to the American Scientist this research has come out after much experimentations on rats.
Experiments proved that mice on a lower calorie diet had comparatively less ability to fight the flu virus.	It has been proven from experiments that rats put on diet with less calories had less ability to resist the Flu virus.	It was proved by experiments the low calories eaters mouses had low defending power for flue in ratio.	Experimentaions have proved that those rats on less calories diet have developed a tendency of not overcoming the flu virus.
research has proven this old myth wrong that its better to fast during fever.	Research disproved the old axiom that " It is better to fast during fever"	The research proved this old talk that decrease eating is useful in fever.	This Research has proved the very old saying wrong that it is good to starve while in fever.

Avoiding dieting to prevent from flu	abstention from dieting in order to avoid Flu	Abstain from decrease eating in order to escape from flue	In order to be safer from flu quit dieting
--------------------------------------	---	---	--

Avoid		dieting		to		prevent	from	flu
Abstention	from	dieting	in order	to		avoid		Flu
Abstain		decrease eating	in order	to		escape		flue
quit		dieting		to	be	safer	from	flu

Avoiding dieting to prevent from flu	abstention from dieting in order to avoid Flu	Abstain from decrease eating in order to escape from flue	In order to be safer from flu quit dieting
--------------------------------------	---	---	--

Avoid		dieting		to		prevent	from	flu
Abstention	from	dieting	in order	to		avoid		Flu
Abstain		decrease eating	in order	to		escape	from	flue
quit		dieting		to	be	safer	from	flu

Translate Urdu into English

Help us translate Urdu articles into English. Your translations will be distributed with a [Creative Commons license](#), so that other people can re-use it. This HIT is for people who speak both Urdu and English. Please **do not use** translation software or online machine translation systems like Google translate. Please make sure that your English translation.

- Does not add or delete any information from the original text
- Has the same meaning and style as the original
- Does not contain any spelling errors
- Is grammatical, natural-sounding English

First, please answer these questions about your language abilities:

Is Urdu your native language? Yes No

How many years have you spoken Urdu? years

Is English your native language? Yes No

How many years have you spoken English? years

افغانستان ایشیاء کا ایک ملک ہے جس کا سرکاری نام اسلامی
جموری افغانستان ہے۔

اس کے جنوب اور مشرق میں پاکستان، مغرب میں ایران، شمال
مشرق میں چین، شمال میں ترکمانستان، ازبکستان اور تاجکستان ہیں۔

اردگرد کے تمام ممالک سے افغانستان کے تاریخی، مذہبی اور
ثقافتی تعلق بہت گمراہے۔

اس کے بیشتر لوگ مسلمان ہیں۔

۔۔۔ ملک بالتدبیب اب ائمہ، بنانیمہ، عدیمہ، تکہو، منگولہو،

Informed Consent Form

Purpose of research study: We are collecting translations to improve translation software and to make Wikipedia content accessible in all languages.

Benefits: Although it will not directly benefit you, this study may benefit society by improving how computers process human languages. This could lead to better translation software, improved web searching, or new user interfaces for computers and mobile devices.

Risks: There are no risks for participating in this study.

Voluntary participation: You may stop participating at any time without penalty by clicking on the "Return HIT" button, or closing your browser window.

We may end your participation if you do not have adequate knowledge of the language, or you are not following the instructions, or your answers significantly deviate from known translations.

Confidentiality: The only identifying information kept about you will be a WorkerID serial number and your IP address. This information may be disclosed to other researchers.

Questions/concerns: You may e-mail questions to the principle investigator, [Chris Callison-Burch](#). If you feel you have been treated unfairly you may contact the Johns Hopkins University [Institutional Review Board](#).

Clicking on the "Accept HIT" button indicates that you understand the information in this consent form. You have not waived any legal rights you otherwise would have as a participant in a research study.

Translation of the first sentence goes here.

Translation of the second sentence goes here.

Quality Control Model

- **Sentence features**
 - Language model probability
 - Ratio of source / target sentence lengths
 - Web n-gram match percentage
 - Translation edit rate to other translators
- **Turker features**
 - Aggregate of sentence feature scores
 - Self-reported language abilities
 - (Is native speaker? How long speaking?)
 - Worker location (Pakistan? India?)
- **Ranking features** (based on second pass vote)
- **Calibration feature** (Bleu against professionals²⁶)

Vote for the best translation

Please read the sentences and vote on the one that you think is the best in each group. The sentences are translations that were produced by people who are not native English speakers. Their translations are often ungrammatical, misspelled, disfluent, or bad in other ways. Your goal is to try to pick the best translation among the set. The one that you choose as the best will be forwarded on for editing, and it will undergo a variety of other quality control mechanisms before it is published.

You should consider the following factors when selecting one translation as the best:

- Does it make more sense than the others?
- Is the English reasonably good?
- Do the grammar and spelling require only minimal correction?

<input type="radio"/>	Experimentaions have proved that those rats on less calories diet have developed a tendency of not overcoming the flu virus .
<input type="radio"/>	in has been proven from experiments that rats put on diet with less calories had less ability to resist the Flu virus.
<input type="radio"/>	Experiments proved that mice on a lower calorie diet had comparatively less ability to fight the flu virus.
<input type="radio"/>	It was proved by experiments the low calories eaters mouses had low defending power for flue in ratio.

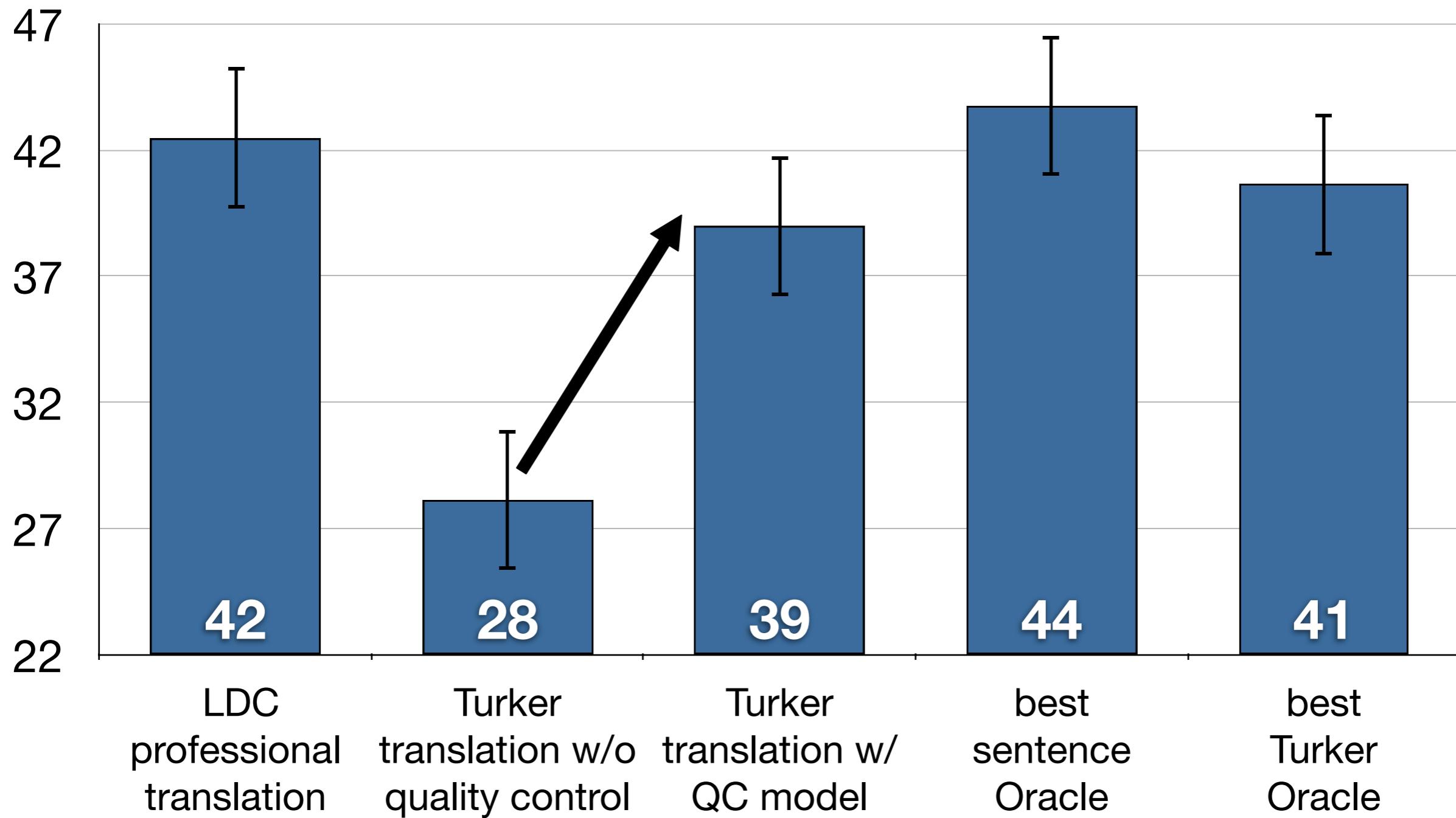
<input type="radio"/>	The research proved this old talk that decrease eating is useful in fever.
<input type="radio"/>	Research disproved the old axiom that " It is better to fast during fever"
<input type="radio"/>	research has proven this old myth wrong that its better to fast during fever.
<input type="radio"/>	This Research has proved the very old saying wrong that it is good to starve while in fever.

<input type="radio"/>	According to the scientist a patient should eat more while in fever.
<input type="radio"/>	According to scientists, eat a lot during fever.
<input type="radio"/>	Eat and drink more in fever according to scientists.
<input type="radio"/>	according to the scientists one should eat a lot during fever.

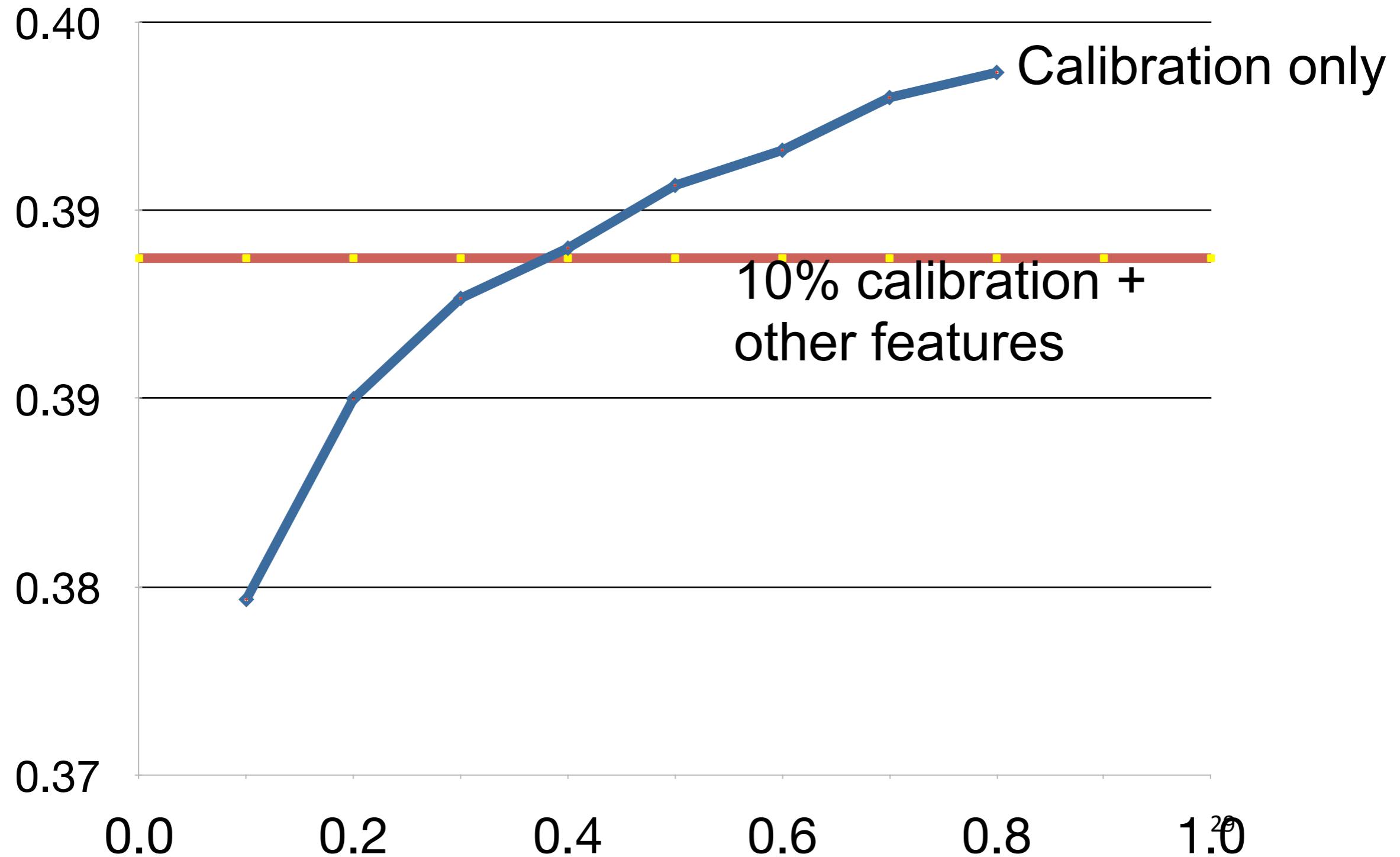
<input type="radio"/>	According to the Researchers of the State of Michigan University diet taking rats have less strength to fight with the infection and also the chances of death increased whereas earlier those rats on common diet had more strength to fight with the flu virus.
-----------------------	---

Professional Quality from Non-Professionals

Full details in Zaidan and Callison-Burch (ACL 2011a) & Zaidan (PhD Thesis 2012)



% of gold standard data used for calibration



Editing HIT

Sri Lanka 's forest region was destroyed by agriculture , ~~wooden works , veticinary feeds , etc . ,~~
forestry animal grazing
several commissions ~~were~~ created to protect the remaining forest region were

Sri Lanka is considered ~~as the~~ bird 's sanctionary place . to be a

For further information please see the article on bird sanctionary rights in Indian Subcontinent
the

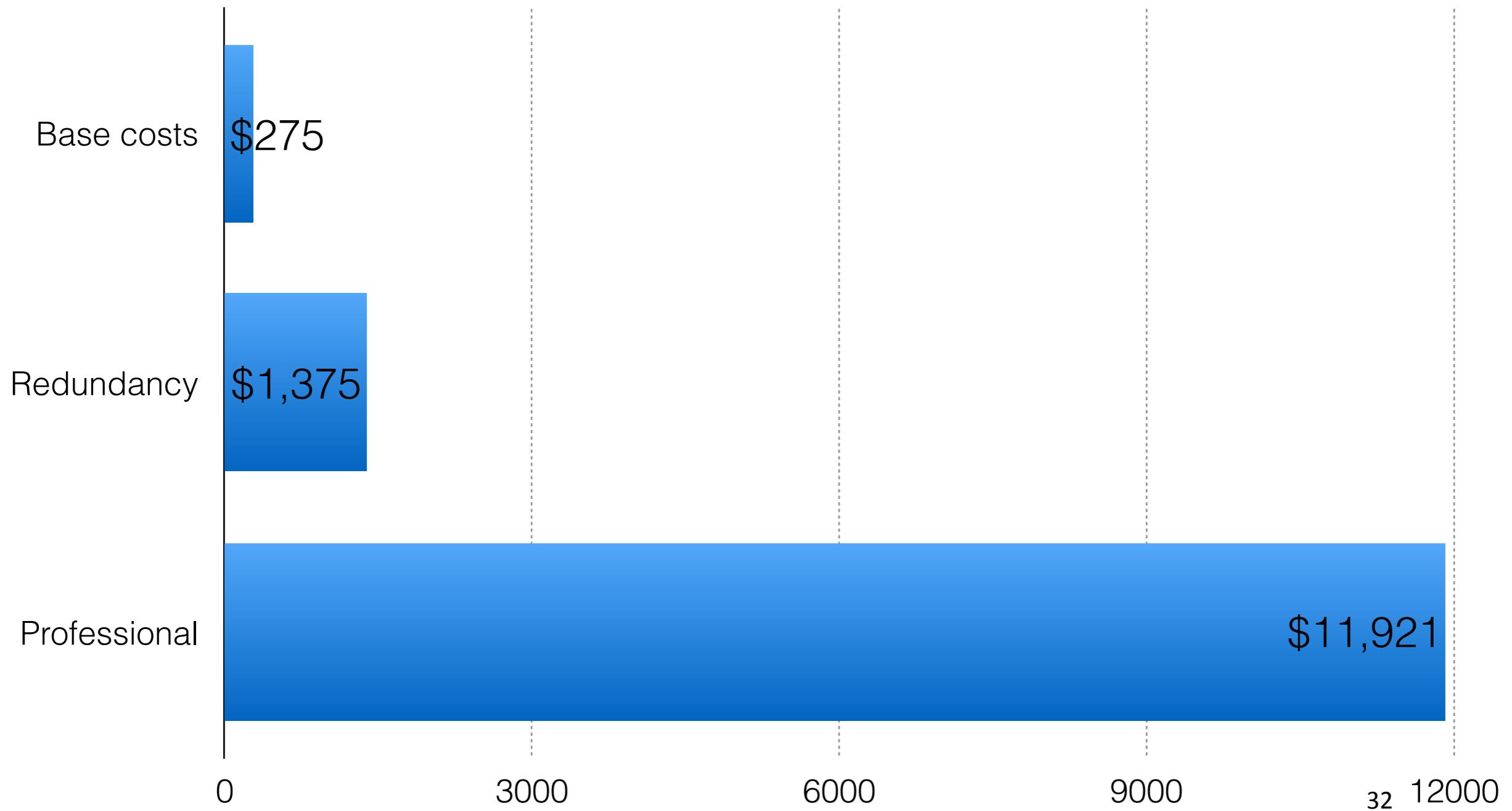
There is thousand of animals living in Sri Lanka which includes several Sri Lanka originated animals .

When we compare the area of Sri lanka 's Island , birds are highly found here .

Cost of Translation on MTurk



Cost of Translation versus Professionals

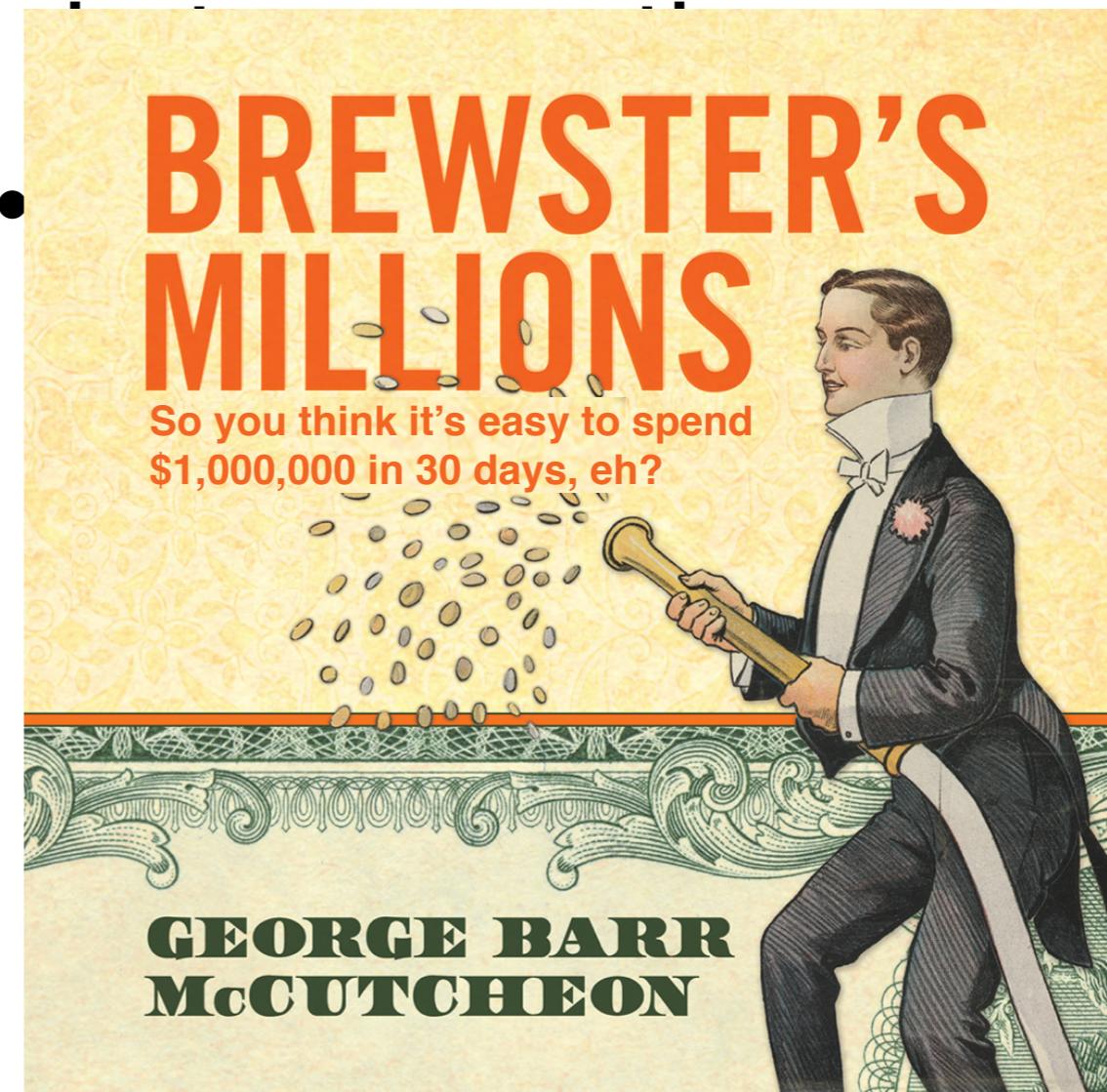


\$100,000 challenge

Surely it can't scale. I bet you can't spend my money fast enough.



- Ken Church bet me \$50k that I couldn't spend \$50k on MTurk



Leaderboard

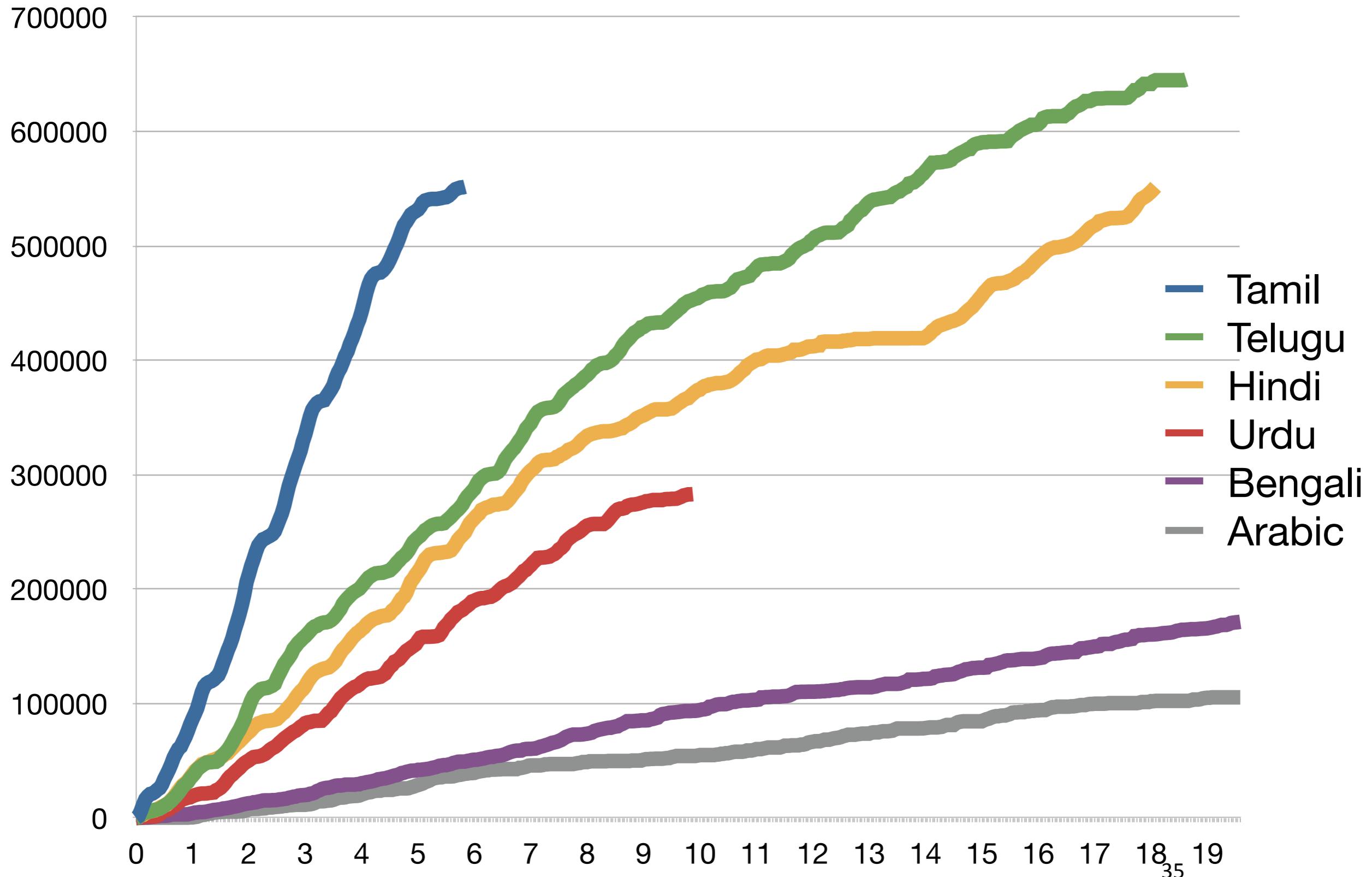
Mechanical Turk Monitor: Top-1000 Recent Requesters

[General](#)[Top requesters](#)[Arrivals](#)[Completed](#)[Search](#)[About](#)

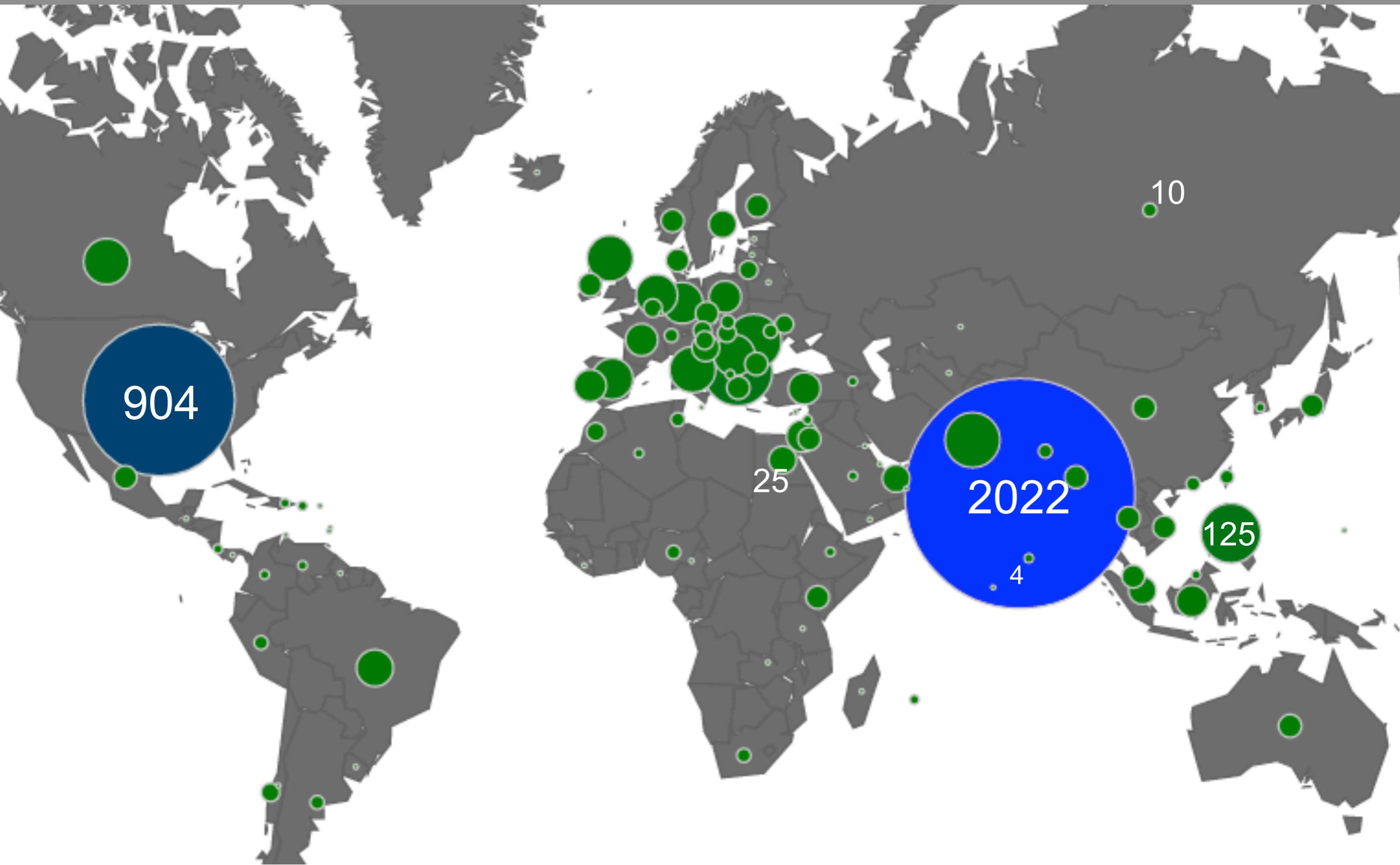
	Requester ID	Requester	#Task	#HITs	Rewards
1	Dolores Labs	A2IR7ETVOIULZU (RSS)	470	317857	33084.24
2	ContentGalore	A2XL3J4NH6JI12 (RSS)	674	10622	17554.6
3	SpeechInk	A1AQ7EJ5P7ME65 (RSS)	9019	13613	12876.67
4	CastingWords	A3MI6MIUNWCR7F (RSS)	9454	14036	8947.02
5	QuestionSwami	AD7C0BZNKYGYV (RSS)	629	4116	4750.37
6	Chris Callison-Burch	A32TTE4XXN6MQZ (RSS)	11	9961	4458.02
7	Smartsheet.com Clients	A1197OGL0WOQ3G (RSS)	434	38212	3118.28
8	retaildata	AD14NALRDOSN9 (RSS)	8	50288	3110.85
9	Classify This	A1CTI3ZAWTR5AZ (RSS)	25	94590	1891.8
10	Andrew Stephen	A1Y25F6MZCMQGY (RSS)	3	22705	1131.25
11	Dolores Labs 2	A3JX8WONBL5N9X (RSS)	34	8976	1043.27
12	RelevanceQuest	A8RMEN71ICE57 (RSS)	15	47881	1029.92
13	Crowd Task	AFAOUHS65HNDS (RSS)	4	2388	955.6
14	nlds.soe.ucsc.edu	A1HI9DWCF794RE (RSS)	4	4702	933.9

Volume Study

Translation Volume: total English words after k days



Language Demographic Study

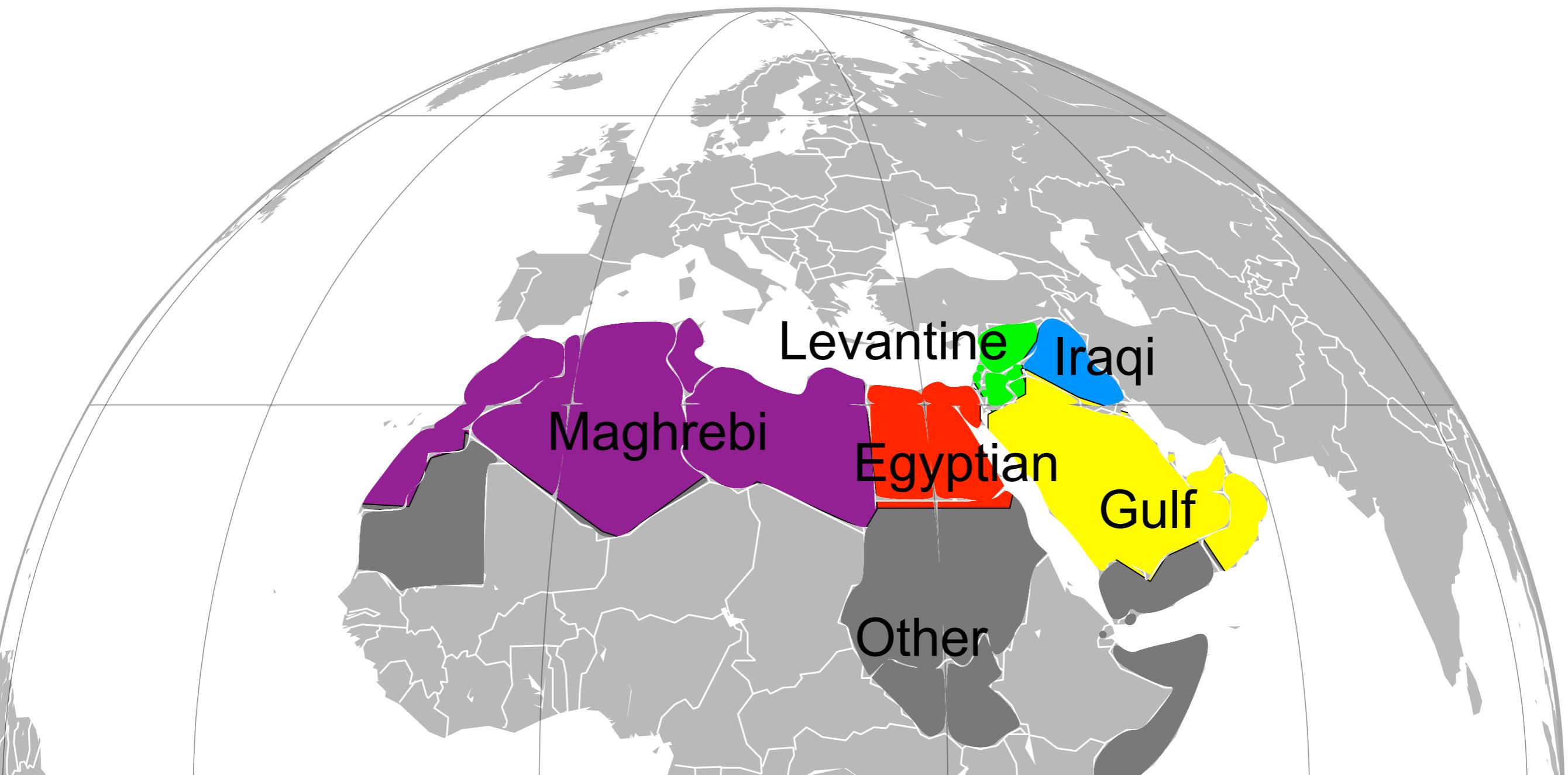


Language Demographic Study

Rank	Language	Native Speakers (millions)	Second Language	Third Language	Fourth Language	Fifth Language
1	English	591	French	63	Vietnamese	34
2	Tamil	253	Polish	61	Macedonian	31
3	Malayalam	219	Urdu	56	Cebuano	29
4	Hindi	149	Tagalog	54	Swedish	26
5	Spanish	131	Marathi	48	Bulgarian	25
6	Telugu	87	Russian	44	Hungarian	23
7	Chinese	86	Italian	43	Swahili	23
8	Romanian	85	Bengali	41	Thai	22
9	Portuguese	82	Gujarati	39	Catalan	22
10	Arabic	74	Hebrew	38	Lithuanian	21
11	Kannada	72	Dutch	37	Punjabi	21
12	German	66	Turkish	35	Others	≤20

Gathering Data about Arabic

Arabic has **different varieties**. MSA is the standardized form but there are many **distinct regional dialects**.



Translating Dialects with MSA MT

MSA:

متى سنرى هذه الثالثة من المجرمين تخضع للمحاكمة ؟

*mtY snrY h*h Alvp mn Almjrmyn txDE llmHAkmp ?*

Levantine:

ايمتى رح نشوف هالشلة من المجرمين بتتحاكم ؟

AvmtY rH n\$wf hAl\$lp mn Almirmvn btthAkkm ?

PT: Quando veremos esse grupo de criminosos serem julgados?

ES-EN: **Quando esse** group of criminals see **Serem julgados**?

Written Arabic Dialect

بلخادم: ما يجمع مصر والجزائر يجعل التقارب أكثر من ضروري

الخميس، 16 يونيو 2011 - 09:27



عبد العزيز بلخادم الممثل الشخصي للرئيس الجزائري

الجزائر (أ.ش.)
+ Google اضف للخرائط

أكد عبد العزيز بلخادم الممثل الشخصي للرئيس الجزائري والأمين العام لحزب جبهة التحرير الوطني ذات الأغلبية في البرلمان والحكومة، أن العلاقات بين مصر والجزائر قديمة ومتعددة علاوة على مصالح مشتركة قوية تربط بين البلدين.

وقال بلخادم إنه بالتأني فإن ما حدث من "سحابات الصيف" "عقب مباراة كرة القدم بين فريقي البلدين في تصفيات كأس العالم عام 2009 لا يؤثر على طبيعة العلاقة والأخوة والروابط بين الشقيقين، مشدداً على ضرورة العمل على رفع سقف التعاون بين البلدين على المستوى الحكومي والجماهيري والمؤسساتي

الملحق، لكنه ألمح إلى إمكانية تشكيل لجنة مشتركة بين البلدين لحل مشكلة تجارة البترول.

Article body (MSA)

خليك في حالك وابعد عننا

بواسطة: bebooo

1

بتاريخ: الخميس، 16 يونيو 2011 - 09:39

EGY

خليك في حالك انت وابعدوا عننا وانت اخر ناس تتكلم عن التقارب والعروبة

الجزائر إخوه أعزاء لنا وأهلاً ومرحباً بتقوية العلاقات معهم

MSA

بواسطة: سامي شاهين

2

بتاريخ: الخميس، 16 يونيو 2011 - 09:55

التقارب بين مصر والجزائر يعتبر حاجة ملحة تستدعيه الظروف الراهنة التي يمر بها الوطن العربي ، أهلاً ومرحباً يأي تقارب عربي بين الأشقاء في الوطن والمصیر ، يسقط الاستعمار الجديد المتمثل في أمريكا والناتو ومن وراءه آل صهيون

هزلت

بواسطة: مواطن مصرى بسيط

3

بتاريخ: الخميس، 16 يونيو 2011 - 09:55

EGY

هزلت وعيلت ومشت سفينة نوح على اليابس - الدنيا جري فيها ايه علشان نسمع دروس من امثالكم

تبليه

بواسطة: علي

4

بتاريخ: الخميس، 16 يونيو 2011 - 09:57

MSA

انا الشعب الجزائري يعمل جاهدا من اجل السعي للتوجيد العالم العربي وكذاك السعي ورا طموحات اكبر وهي ، ان تكون كلمت العرب كلما واحدة امام الغرب

أنت.....

بواسطة: الفيلسوف

5

بتاريخ: الخميس، 16 يونيو 2011 - 10:14

EGY

أنت بتكلم مين يا بلخادم كل الأنظمة فاسدة وتابعة لأمريكا وإسرائيل، إن أردتم فعلا حل مشاكل العرب في البيت العربي فما على الحكم إلا تكوين جمعية وذهب لمتحجع الحكم المخلوعين بالمملكة السعودية وسيروا الشعوب تبحث في قيام الولايات العربية المتحدة وعاصمتها القدس إنشاء الله (معلش حبيبي هو حلم يقطن يمكن يتحقق في يوم من الأيام من هنا يعلم الغيب؟)

MSA

EGY

Crowdsourcing Arabic Dialects

- Translated dialect-labeled segments  

Dialect Classification HIT	\$10,064
Sentence Segmentation HIT	\$1,940
Translation HIT	\$32,061
Total Cost	\$44,065
Num words translated	1,516,856
Cost per word	2.9 cents/word

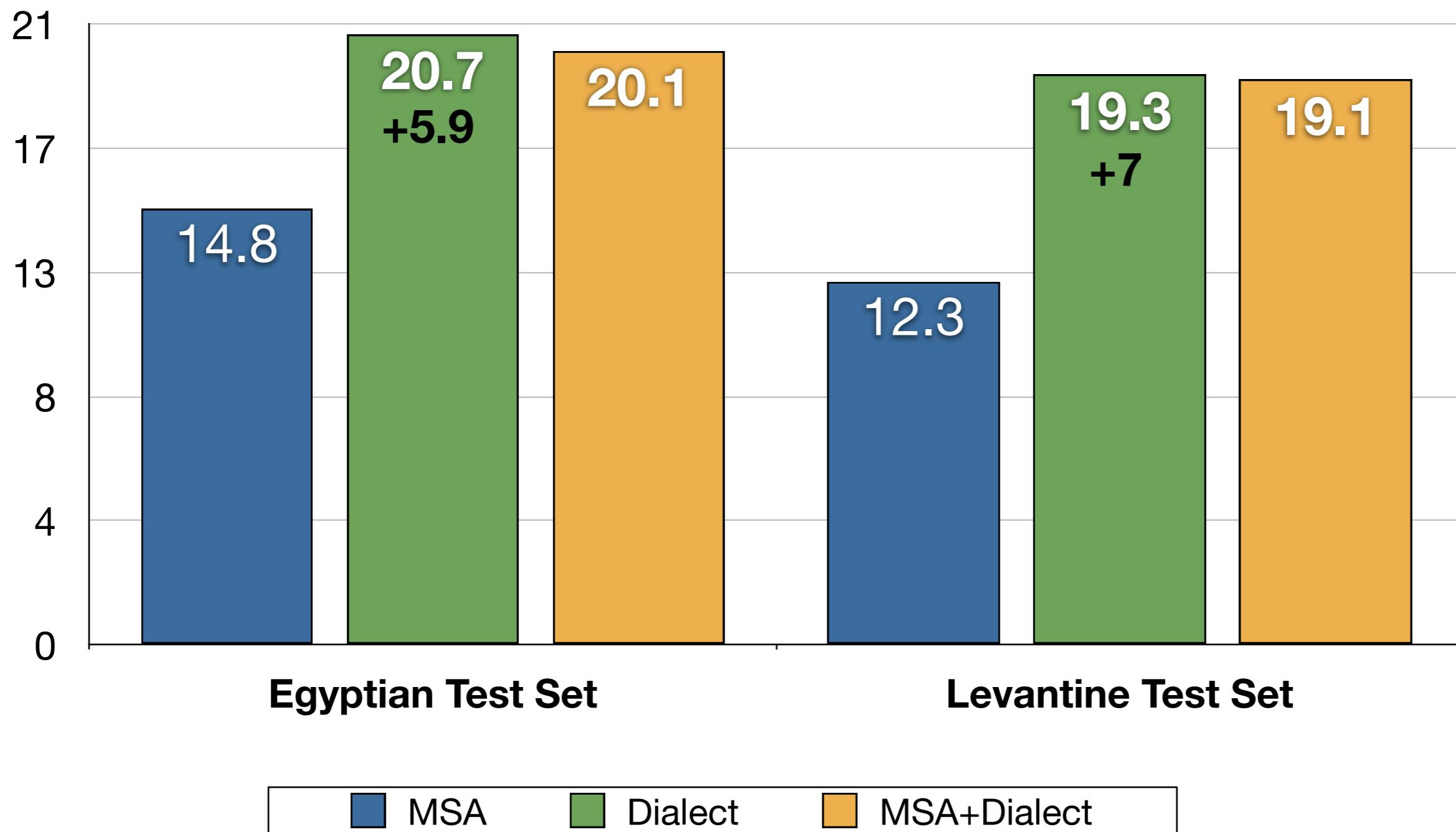
- 121 workers completed 20+ assignments
- 200,000 words translated per week
- Trained BBN's machine translation system

Examples of Dialect Translation

Dialect Input	MSA system	Dialect system	Reference
EGY انت بتعمل له اعلان ولا ايه ؟ !!	You are working for a declaration and not?	You are making the advertisement for him or what?	Are you promoting it or what?!!
EGY نفسي اطمئن عليه بعد ما شاف الصوره دي	Myself feel to see this image.	I wish to check on him after he saw this picture.	I want to be sure that he is fine after he saw the images.
LEV لهيك الجو كتير كوروول	God you the atmosphere.	This is why the weather is so cool	This is why the weather is so cool
LEV طول بالك عم نمزح	Do you think about a joke long.	Calm down we are kidding	Calm down, we are only kidding

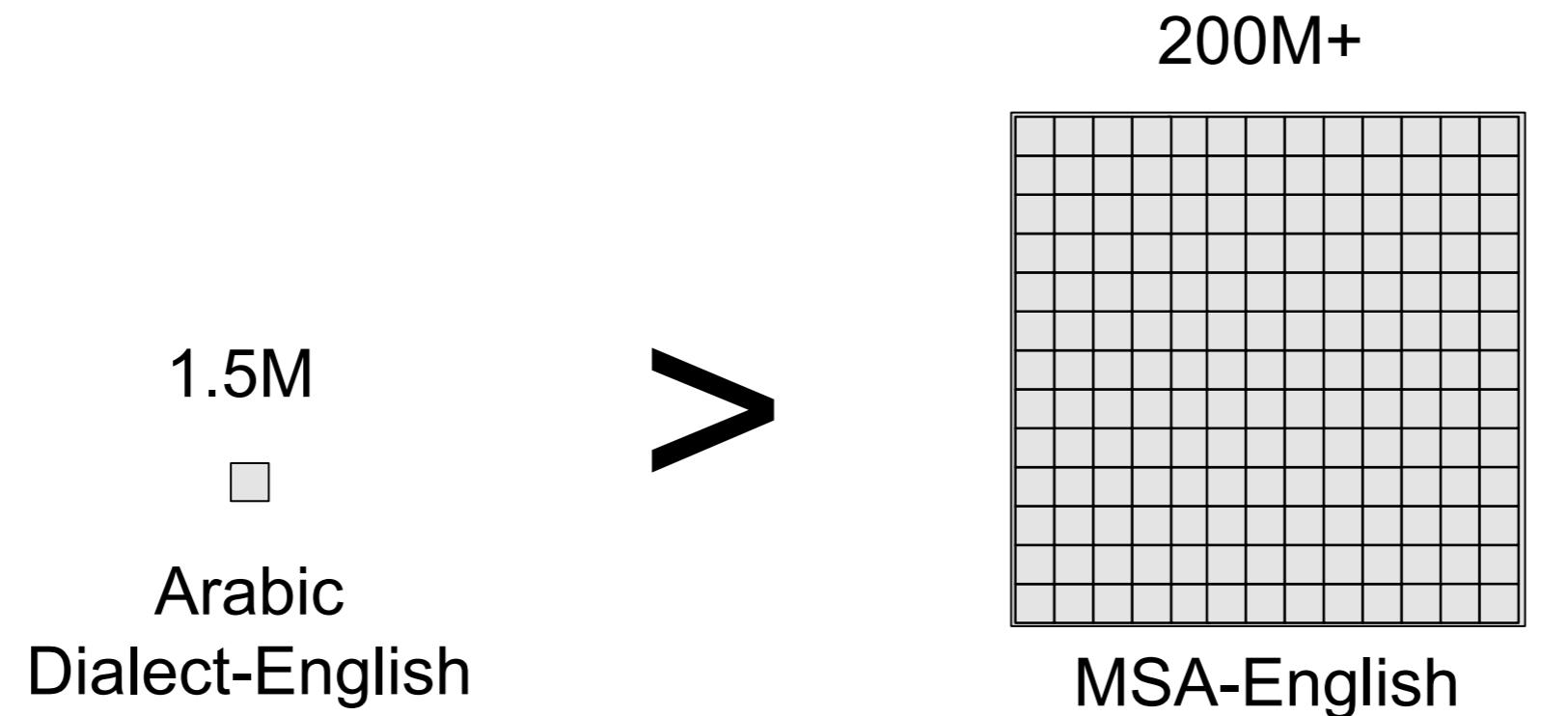
Dialect versus MSA

Full details in Zbib, Malchiodi, Devlin, Stallard, Matsoukas, Schwartz, Makhoul, Zaidan, & Callison-Burch (NAACL 2012)



Dialect v. MSA data for MT

For machine translation of Arabic dialects



2 months

\$44k

+6 Bleu

5 years

\$10s of millions

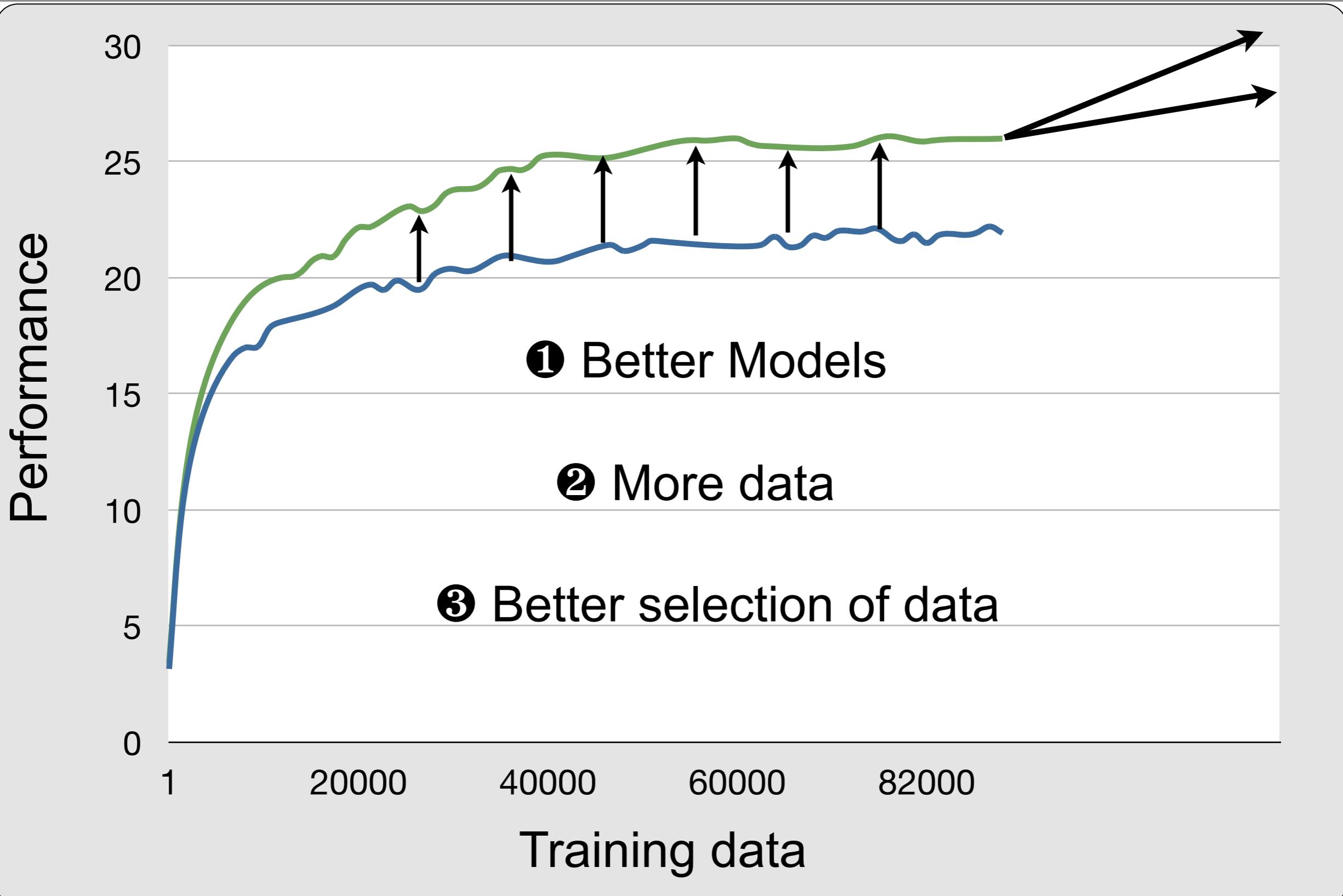
no added Bleu

Implications of low cost, high volume data annotation for research

Individual researchers now have their own data production companies

- ✓ Parallel corpora for six Indian languages
 - ✓ Arabic dialect ID dataset
 - ✓ 1.5M word English-Arabic dialect corpus
 - ✓ Bilingual dictionaries for ~100 languages
- ✓ English translation of 180 hrs of spoken Spanish
- ✓ Speech collection and transcription for new lang.
- ✓ De-romanization / text normalization corpus
- ✓ 10s of thousands of paraphrase judgments
- ✓ Tons of data for natural language understanding

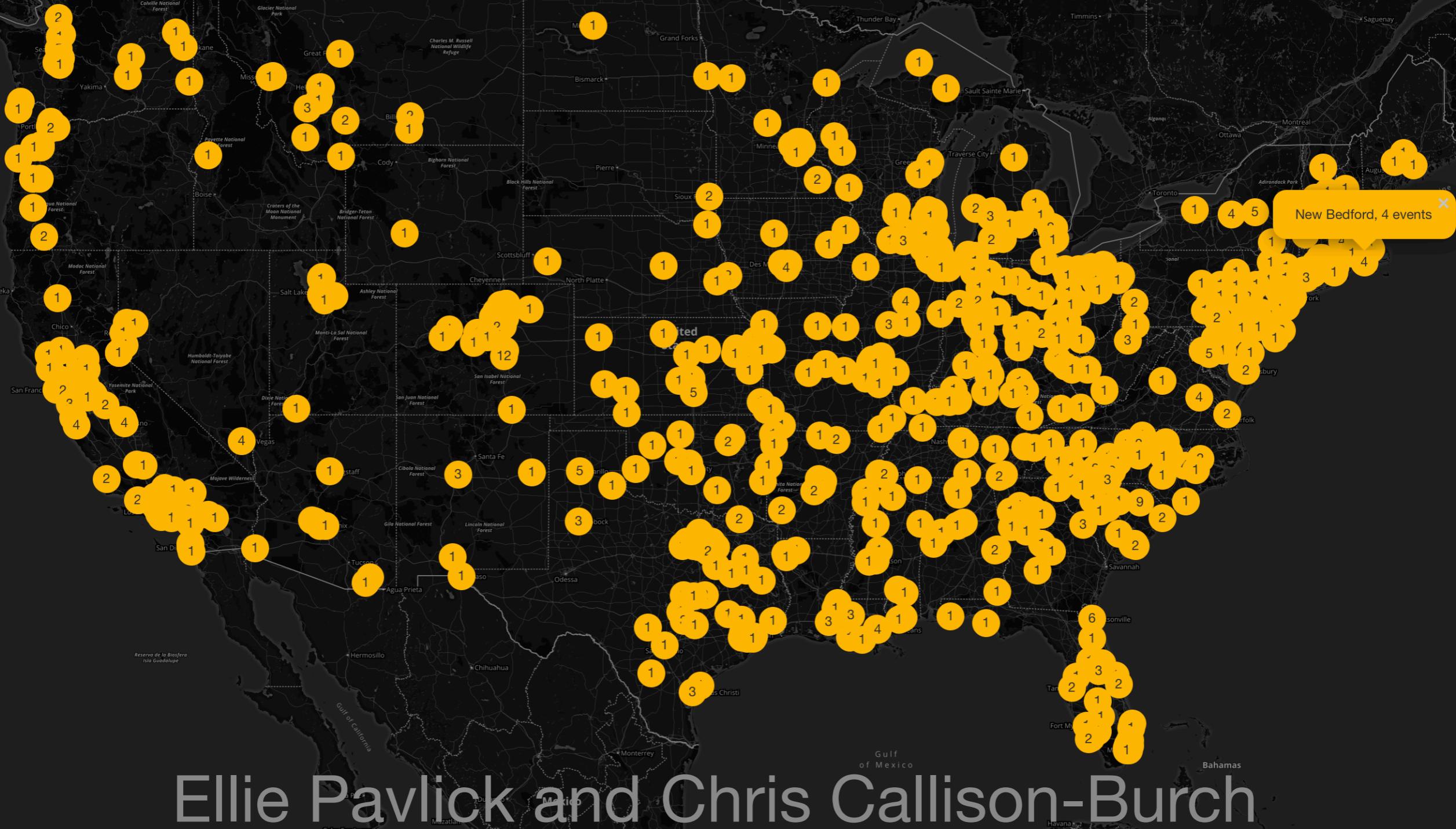
Crowdsourcing and Machine Learning



Crowdsourcing and you

- You can conduct large-scale human studies to
 - validate theories from social science
 - do large-scale machine learning experiments
 - build next generation HCI interfaces
 - study economic markets or political trends
- Crowdsourcing provides low-cost, instant-on access to a large crowd of workers
- Try it out yourself!
- I'm happy to provide advice on how to design good Mechanical Turk tasks

The Gun Violence Database



Ellie Pavlick and Chris Callison-Burch
University of Pennsylvania

Goals of the GVDB

Collect data about gun violence in the US
to facilitate **public health research**

Draw sample from local newspapers and
television stations that publish online

Use machine learning and crowdsourcing
to **extract structured data from text**



An engraving of the Mechanical Turk, the 18th century chess-playing automaton

Crowdsourcing and human computation are emerging fields that sit squarely at the intersection of economics and computer science. They examine how people can be used to solve complex tasks that are currently beyond the capabilities of artificial intelligence algorithms. Online marketplaces like [Mechanical Turk](#) and [CrowdFlower](#) provide an infrastructure that allows micropayments to be given to people in return for completing human intelligence tasks. This opens up previously unthinkable possibilities like people being used as function calls in software. We will investigate how crowdsourcing can be used for computer science applications like machine learning, next-generation interfaces, and data mining. Beyond these computer science aspects, we will also delve into topics like the sharing economy, prediction markets, how businesses can capitalize on collective intelligence, and the fundamental principles that underlie democracy and other group decision-making processes.

Course number

[NETS](#) 213 - students from all majors are welcome!

Instructors

[Chris Callison-Burch](#) and [Ellie Pavlick](#)

Teaching Assistants

[Course Staff](#)

Discussion Forum

[Piazza](#)

<http://crowdsourcing-class.org>

Time and place

Spring 2016, MWF 2-3PM, LRSM Auditorium

Office Hours

[See calendar page](#)

Prerequisites

[CIS 120](#) or prior programming experience

Course Readings

Each lecture has an accompanying set of [academic papers](#)

Grading

This is a project-based course. Instead of exams, you will do a series of hands-on assignments and a final project.

- Weekly assignments (45%)
- Final project (45%)
- Peer grading (5%)
- Participation (5%)



THE NEW OLD AGE
As Their Numbers Grow,
Home Care Aides Are
Stuck at \$10.11



Will Women Play Major
League Baseball? (And Not
Just on TV)



WELL
Too Old to Donate Blood?



HEALTH | VITAL SIGNS: CAUSE AND EFFECT

VITAL SIGNS: CAUSE AND EFFECT; *Linking Guns and Gun Violence*

By ERIC NAGOURNEY MAY 27, 2003



People with guns in their homes are almost twice as likely to be killed by guns as people who do not keep them at home, researchers reported yesterday in *The Annals of Emergency Medicine*.

And, the researchers found, people with guns are 16 times as likely to commit suicide using guns.

The explanation may lie in the unforgiving nature of firearms, said the author of the study, Dr. Douglas J. Wiebe, who conducted the research at the University of California at Los Angeles and is now at the University of Pennsylvania.

"People who are shot are substantially more likely to die than people injured with nongun weapons," Dr. Wiebe said.

The study was based on a review of the deaths of 1,720 homicide victims and 1,959 suicide victims and a sampling of American adults.

It found that most of the victims, over 56 percent, knew their assailants. A fifth of the homicides occurred during robberies, 6 percent during drug deals and about 15 percent during family arguments.

The study also found that women were significantly more likely than men to be victims of gun homicides. "This likely reflects the singular danger faced by women in abusive relationships," Dr. Wiebe wrote.



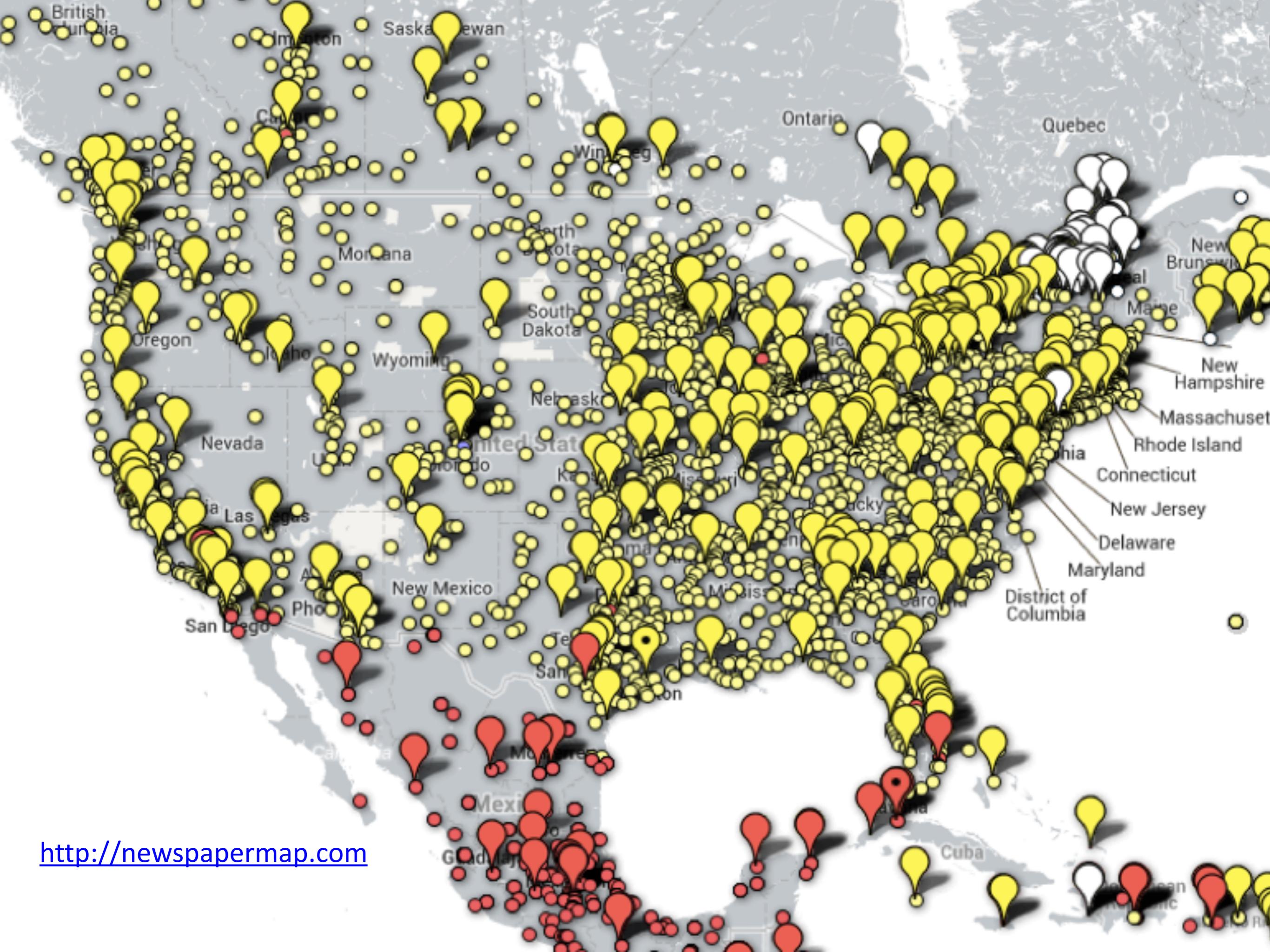
Douglas Wiebe
Professor of Epidemiology
University of Pennsylvania
Perelman School of Medicine

Congressman Who Restricted Gun Violence Research Has Regrets

“Compared to five years ago, the funding picture for a few of us who have done this work for a long time is rosy,” Wintemute said. “Compared to what it requires, it is still bleak. We have lost 20 years of concentrated effort.”

Others have found the field fairly difficult to traverse. Dr. Douglas Wiebe, an associate professor of epidemiology at the Perelman School of Medicine at the University of Pennsylvania, worked on [a 2009 study](#) on the link between gun possession and gun assault that is believed to have sparked Congress’ interest in applying the Dickey amendment to the NIH. He called the restriction of funds “not fatal” to his field, “but very close to it.” Investigators, he explained, are being forced toward less-politically contentious studies, which makes it close to impossible to conduct sound epidemiological research.







Woman stabbed in Ogden incident released from hospital

Yes

No

Unclear

[Show full text](#)

Wrangler News - Online Edition - Home

Yes

No

Unclear

Humans manually verify the predictions of the classifier.

Probe underway after man shot as marshals served warrant

Yes

No

Unclear

[Show full text](#)

POLICE KILL CAR THEFT SUSPECT

Yes

No

Unclear

[Show full text](#)



Woman stabbed in Ogden incident released from hospital

Show full text

Yes

No

Unclear

Wrangler News

Show full text

Unclear

Unclear

LIVE Budget clash in Rajya Sabha

Show full text

Unclear

Wrangler News - Online Edition - Home

<http://www.wranglernews.com/053108.htm>

If your kids are enrolled in the Hoops Star camp this summer at Kiwanis Park, you can expect they'll be learning basketball fundamentals and shooting skills from a pair of coaches who really know the game, namely the two Sam Duanes, senior and junior. [Calendar](#) | [Classifieds](#) | [Contact Us](#) | [Home](#) | [Make a Payment](#) | [Media Kit](#) | [Online Advertising](#) | [Online Map](#) | [Online Pages](#) | [Previous Issues](#) | [Submit Your Ad](#) Copyright ? 2008 Wrangler News

Probe underway after man shot as marshals served warrant

Show full text

Yes

No

Unclear

POLICE KILL CAR THEFT SUSPECT

Show full text

Yes

No

Unclear

PUBLICATION DATE: AUGUST 1, 2016

Police: Officer shoots, wounds shoplifting suspect outside Conroe Wal-Mart - Houston Chronicle

Conroe police say an officer shot and wounded a suspected shoplifter in the parking lot of a Wal-Mart store Monday – the second officer-related shooting of someone suspected of taking merchandise from a city Wal-Mart store in less than three years. According to investigators, Fillmore was observed concealing several items of merchandise and then leaving the store when a Wal-Mart employee attempted to stop him. A Conroe police officer who responded to the incident feared for the safety of the Wal-Mart employee and other citizens in the busy parking lot and fired one shot, wounding the suspect in his left shoulder. Fillmore has convictions dating to 1977 for various offenses including home burglaries, thefts and illegal drug possession, public records show.

First answer a series of binary questions about the circumstances of the shooting....

Please read the text carefully, and then select an answer for all questions. Please base your answers only on information that is explicitly stated or can be confidently inferred from the text of the article.

The shooting was unintentional.

 Yes No Not Mentioned

The shooting was by a police officer.

 Yes No Not Mentioned

The shooting was directed at a police officer.

 Yes No Not Mentioned

The firearm was stolen.

 Yes No Not Mentioned



PUBLICATION DATE: JULY 26, 2016

Man shot in North Baltimore and checks himself into hospital

A 32-year-old man was shot in North **Baltimore** and checked himself into a hospital Wednesday **evening**, police said. Detectives determined the man was shot on the 4600 block of Midwood Ave. in the Winston-Govans neighborhood, police said.

Officers had been called to the scene. Police did not release the man's condition. Anyone with information can call 410-396-2221.

[Clear this highlight](#)[Clock Time \(1p.m., 2:37a.m\)](#)[Additional Location Details](#)

First, try to figure out the date of the described event, and select it by clicking on the calendar icon. The publication date and the day of week mentioned in article are helpful in determining the date of the shooting. Next, click on parts of the text that correspond to the other information listed below, if that information is present in the article. When you highlight a passage of text in the article, you will get a dropdown menu that lets you select which question it answers.

Date

2016-07-20



State

MD - Maryland



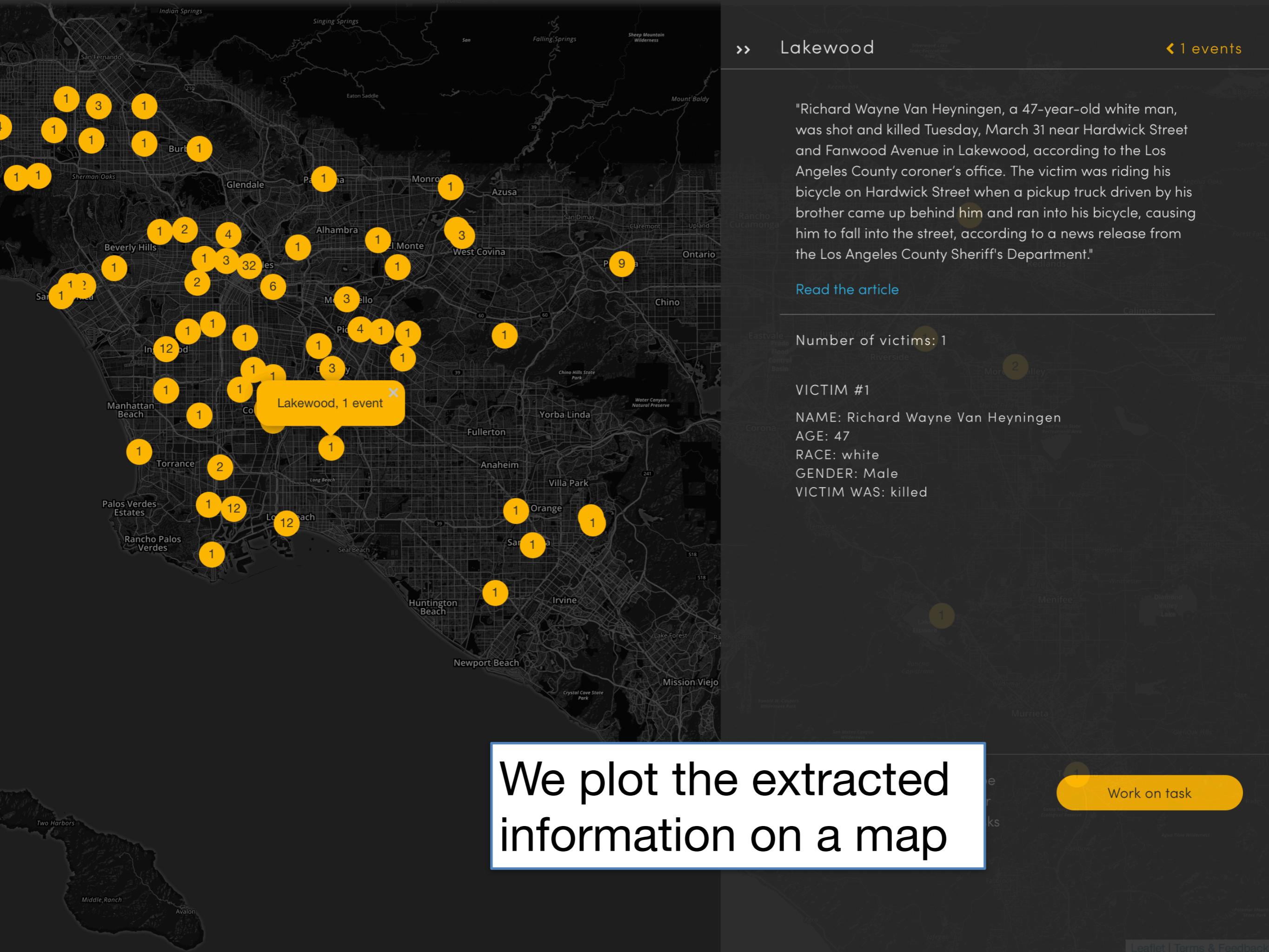
City

Baltimore



Clock Time (1p.m., 2:37a.m)

...then extract structured information from text to populate the database.



Ethical questions about Mechanical Turk

Amazon Mechanical Turk: Gold Mine or Coal Mine?

Karën Fort*
INIST-CNRS/LIPN

K. Bretonnel Cohen†
U. Colorado School of Medicine/U.
Colorado at Boulder

Gilles Adda**
LIMSI/CNRS



Recently heard at a tutorial in our field: “It cost me less than one hundred bucks to annotate this using Amazon Mechanical Turk!” Assertions like this are increasingly common, but we believe they should not be stated so proudly; they ignore the ethical consequences of using MTurk (Amazon Mechanical Turk) as a source of labour.

Here's an excerpt from an IRB application Chris Callison-Burch posted - "We will pay participants small sums of money to complete our tasks, ranging from \$0.01 to \$1. All participants can choose for themselves whether the compensation is fair, and opt not to do it if they deem the compensation to be too low. Amazon's Mechanical Turk has many other researchers and companies offering tasks, so we will offer compensation that is similar to what others offer."

He first refers to Amazon's Mechanical Turk as "an online labor market." And that, I agree with. It is an online labor market.

Requesters like him, and CrowdFlower, **collude**, explicitly or implicitly, **to keep wages at a substandard level** that is compatible with existence standard.

Unlike Jewels, I don't blame workers for taking low paying jobs. I can't blame a person for being needy enough to take what amounts to a crust of bread. **I blame Chris Callison-Burch, and others like him, for keeping the standard wage at crust of bread level.**

I feel like Maria in "Metropolis."

I tried one of those to see, I gave it up at 4 minutes in and about 2/3 of the way through. For the whole hit, I'd have taken about 6 minutes. 10 hits an hour - **\$1.70 an hour**. Restricted to U.S. residents.

This is far too low to be considered a fair wage for a U.S. resident. My performance may be very far off from what others can do. Perhaps I took 4 times or more as long as an average worker would.

My complaint is that any U.S. requester knows what wage rate is required for a U.S. resident to survive. We may not agree on an exact number. But as they say, I know a fair wage when I see it, and this is not it.

Mturk is actually much smaller than what it can appear to be. Something close to requester monopoly has the power to keep wages low. Requester co-operation, explicit or implicit, reinforces this.

Chris Callison-Burch is not unaware, I think, of the mechanics of the wage structure of Mturk.

**WORKERS
OF THE WORLD
UNITE!**



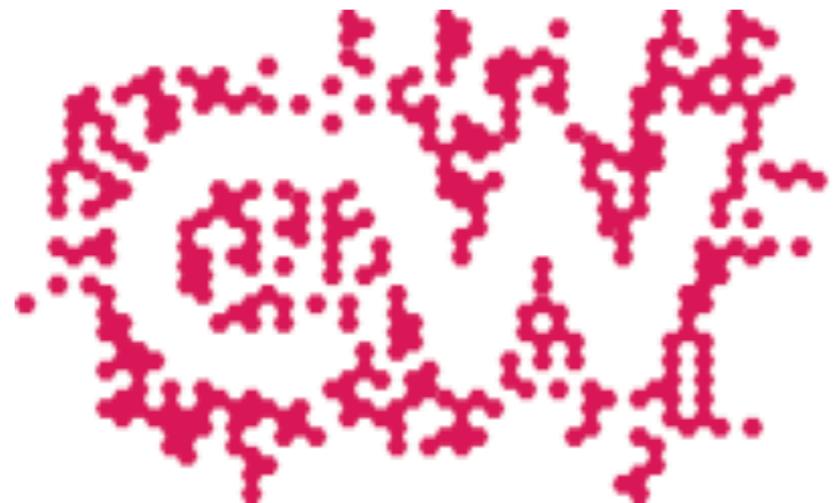
Web Workers Unite! Addressing Challenges of Online Laborers

Bederson and Quinn (ALT CHI 2011)

The ongoing rise of human computation has created an environment where human workers are often regarded as nameless, faceless computational resources. Some people have begun to think of online tasks as a “remote person call”.

Unfortunately, as with any labor market, once humans and money are involved, a host of problems surface. From privacy breaches to unpaid or underpaid labor, there are real social risks that arise with the use of these technologies

It is our responsibility to address them since it is designers, not ethicists or policy makers, who have the power to influence what is built and to mitigate risks before any harm is done.



Crowd Workers

I developed a browser plug-in to improve Mechanical Turk for Workers

It tracks and aggregates statistics across many workers, so that they can find higher paying work

Academically, I am interested in these questions

- a) How much time does the average Turker spend working?
- b) What is their hourly rate?
- c) How reputable are requesters (what fraction of the HITs do they approve v. reject)?
- d) How much time to workers spend searching v working?



Crowd
Workers

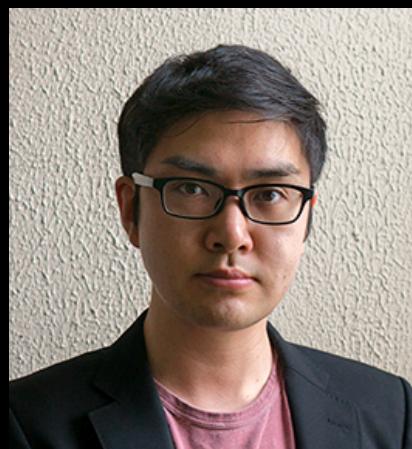
3.8m
task records

2.7k
workers

A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk CHI 2018

Worker Demographics and Earnings on Amazon Mechanical Turk: An Exploratory Analysis CHI 2019

The Gender Wage Gap in an Online Labour Market: The Cost of Interruptions In Submission



Kotaro Hara



Abigail Adams



Kristy Milland



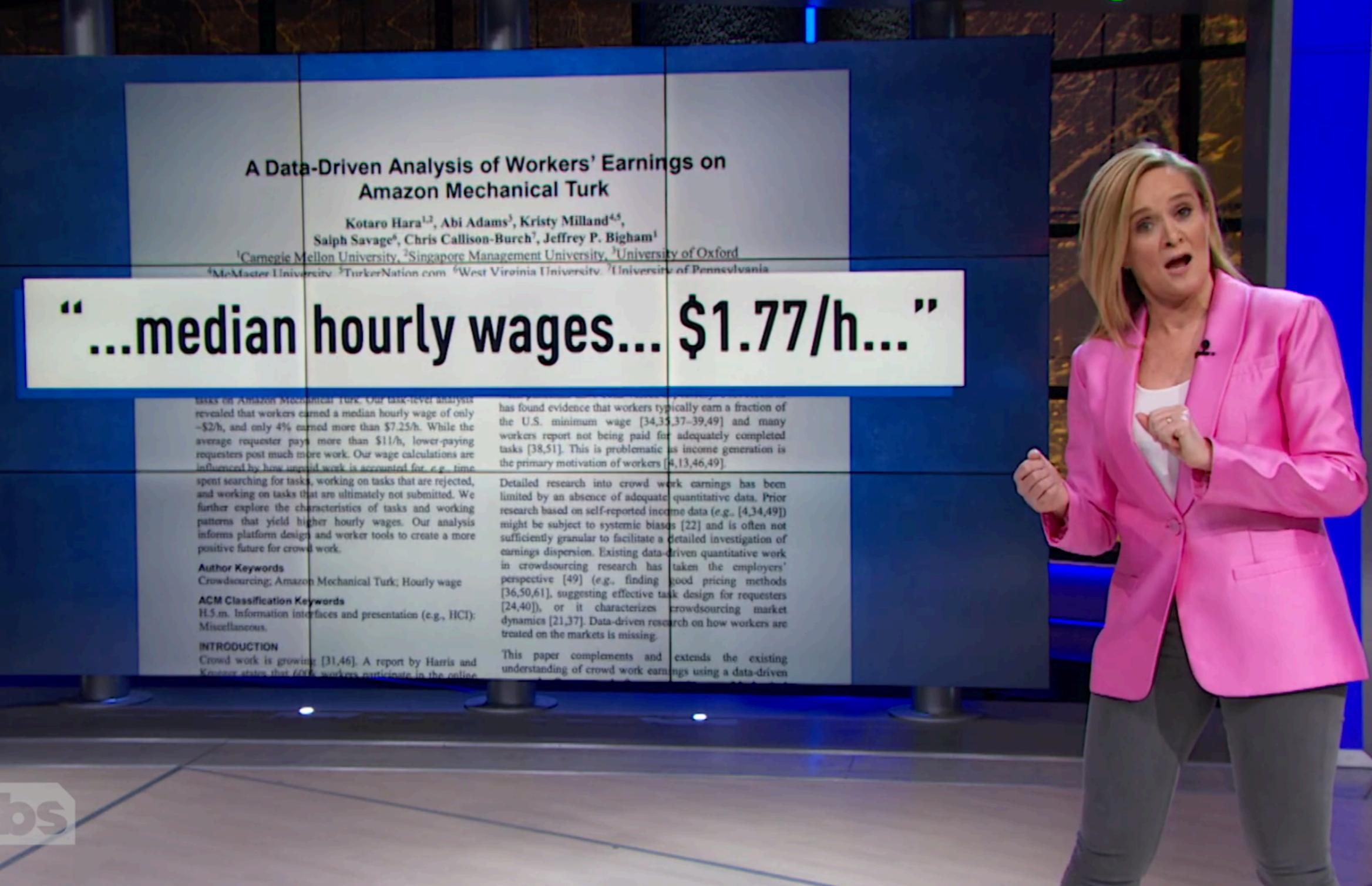
Saiph Savage



Chris Callison-Burch



Jeffrey P. Bigham



Full Frontal with Samantha Bee - March 11, 2020 - America's
Gig-Based Economy Gets Zero Stars

The Gender Wage Gap in an Online Labour Market: The Cost of Interruptions In Submission

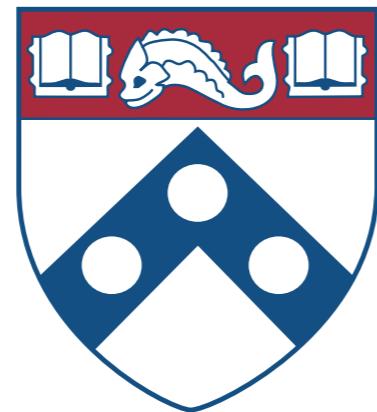
Women earn 20% less per hour on average on Amazon Mechanical Turk. This isn't explained by task selection or experience on the platform. Half of this gap is explained by differences in the scheduling of work; women have more fragmented work patterns with consequences for their task completion speed. A follow up survey shows that the wage gap is concentrated amongst women with young children.



Abigail Adams-Prassl

Thanks!

Chris Callison-Burch
ccb@upenn.edu



Penn
UNIVERSITY *of* PENNSYLVANIA

NP1
hamd ansary

NP2
na}b sdr

P3
kylye

V4
namzd

AUX5
taa

NP1
Hamid Ansari

NP2
Vice President

P3
for

V4
nominated

AUX5
was

