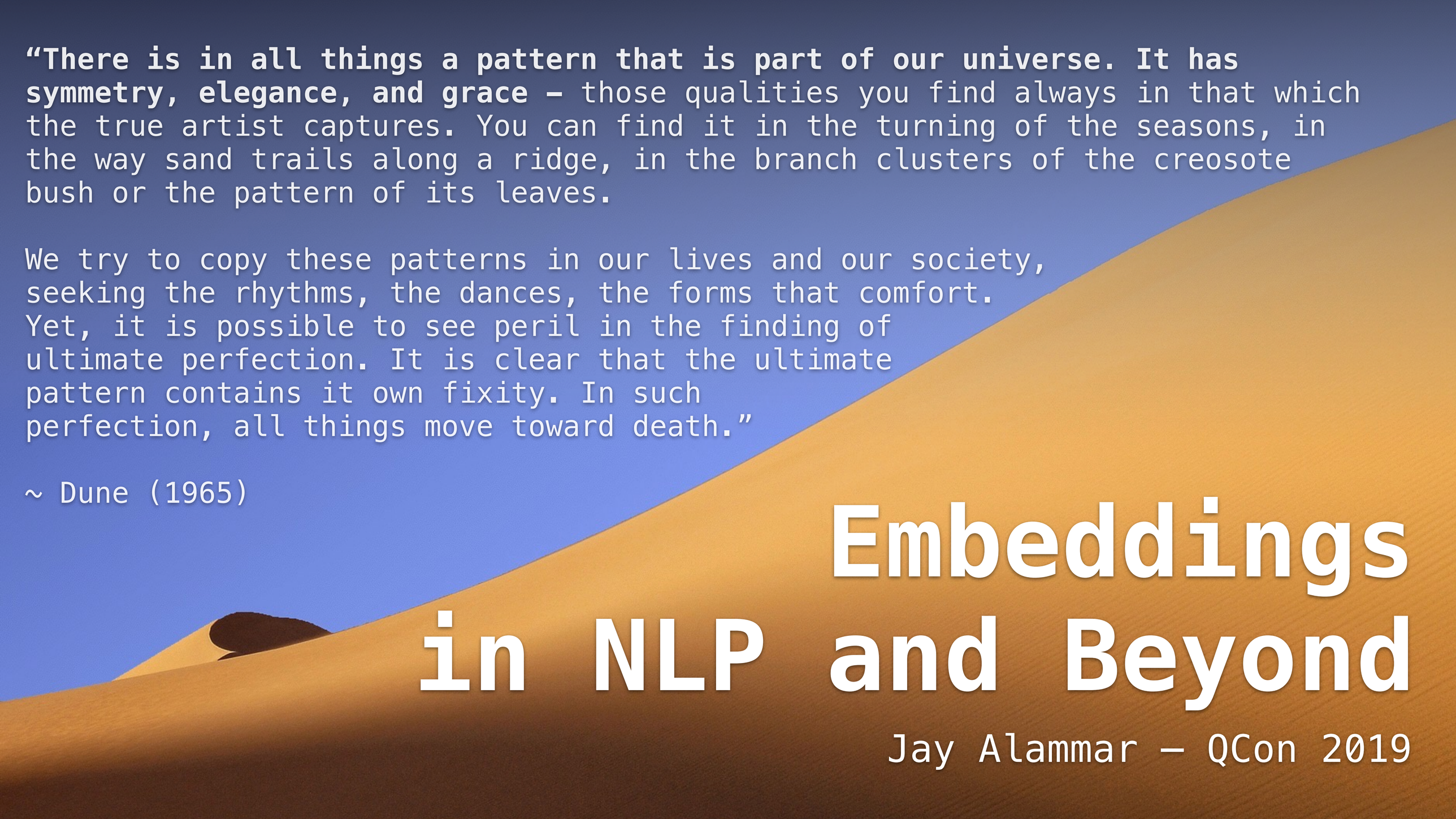
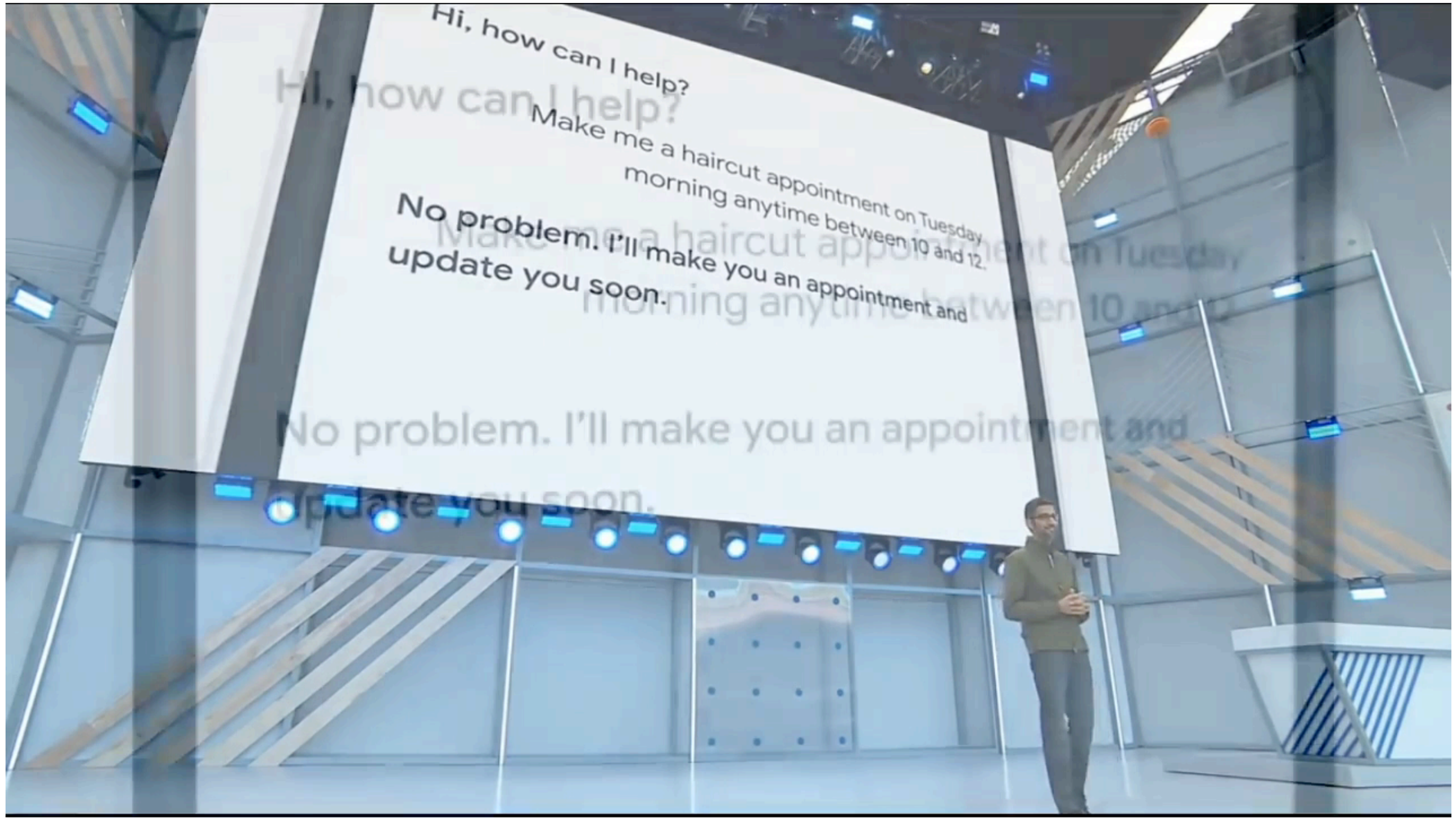
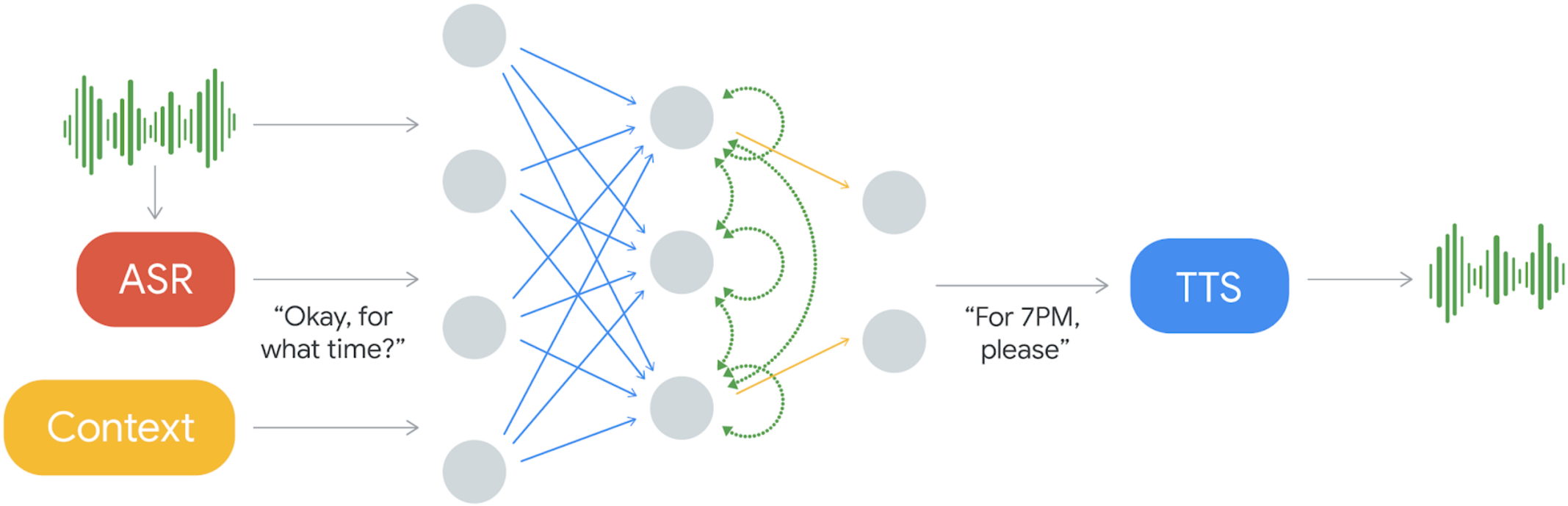

“There is in all things a pattern that is part of our universe. It has symmetry, elegance, and grace – those qualities you find always in that which the true artist captures. You can find it in the turning of the seasons, in the way sand trails along a ridge, in the branch clusters of the creosote bush or the pattern of its leaves.
We try to copy these patterns in our lives and our society, seeking the rhythms, the dances, the forms that comfort. Yet, it is possible to see peril in the finding of ultimate perfection. It is clear that the ultimate pattern contains it own fixity. In such perfection, all things move toward death.”
~ Dune (1965)
Embeddings in NLP and Beyond
Jay Alammar – QCon 2019



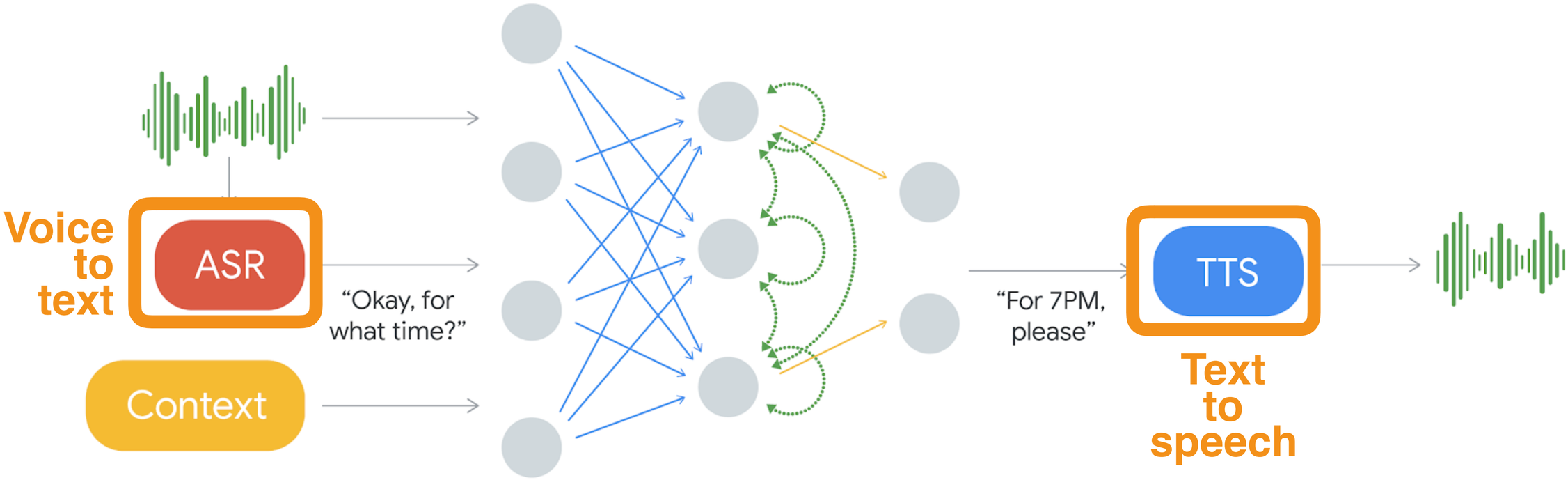
Google Duplex: A.I. Assistant Calls Local Businesses To Make Appointments

https://www.youtube.com/watch?v=D5VN56jQMWM

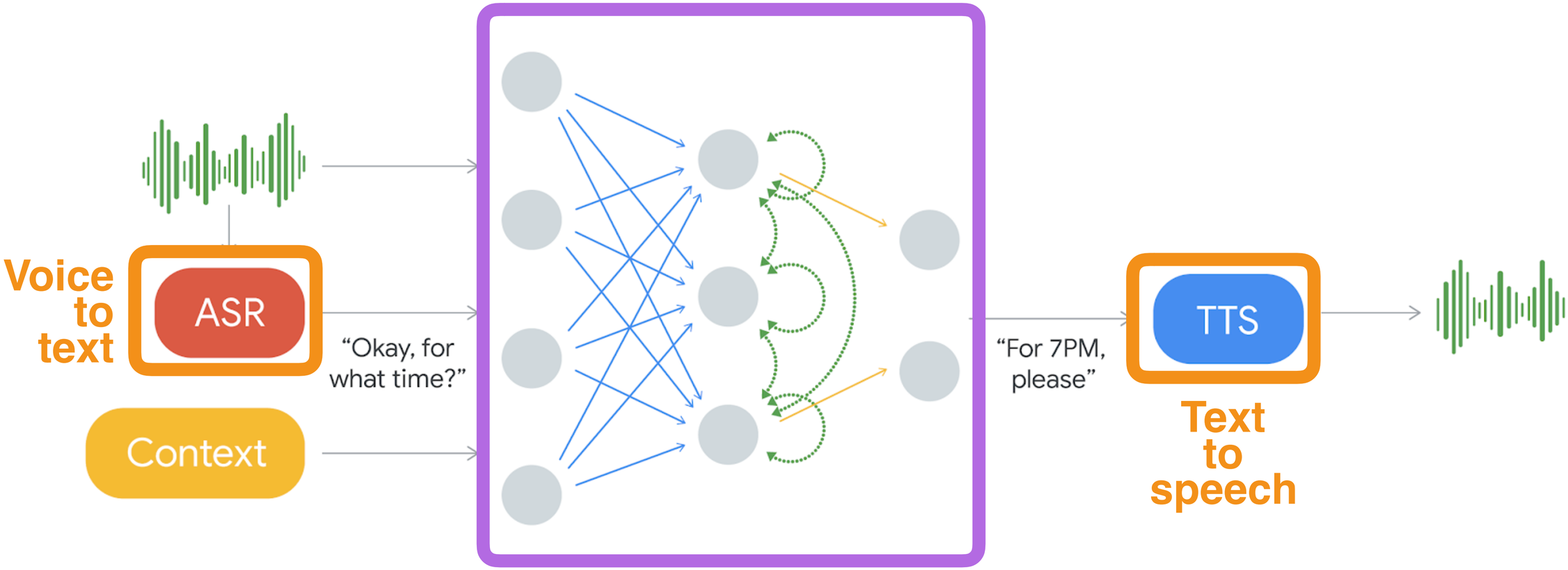
0:45 - 02:14



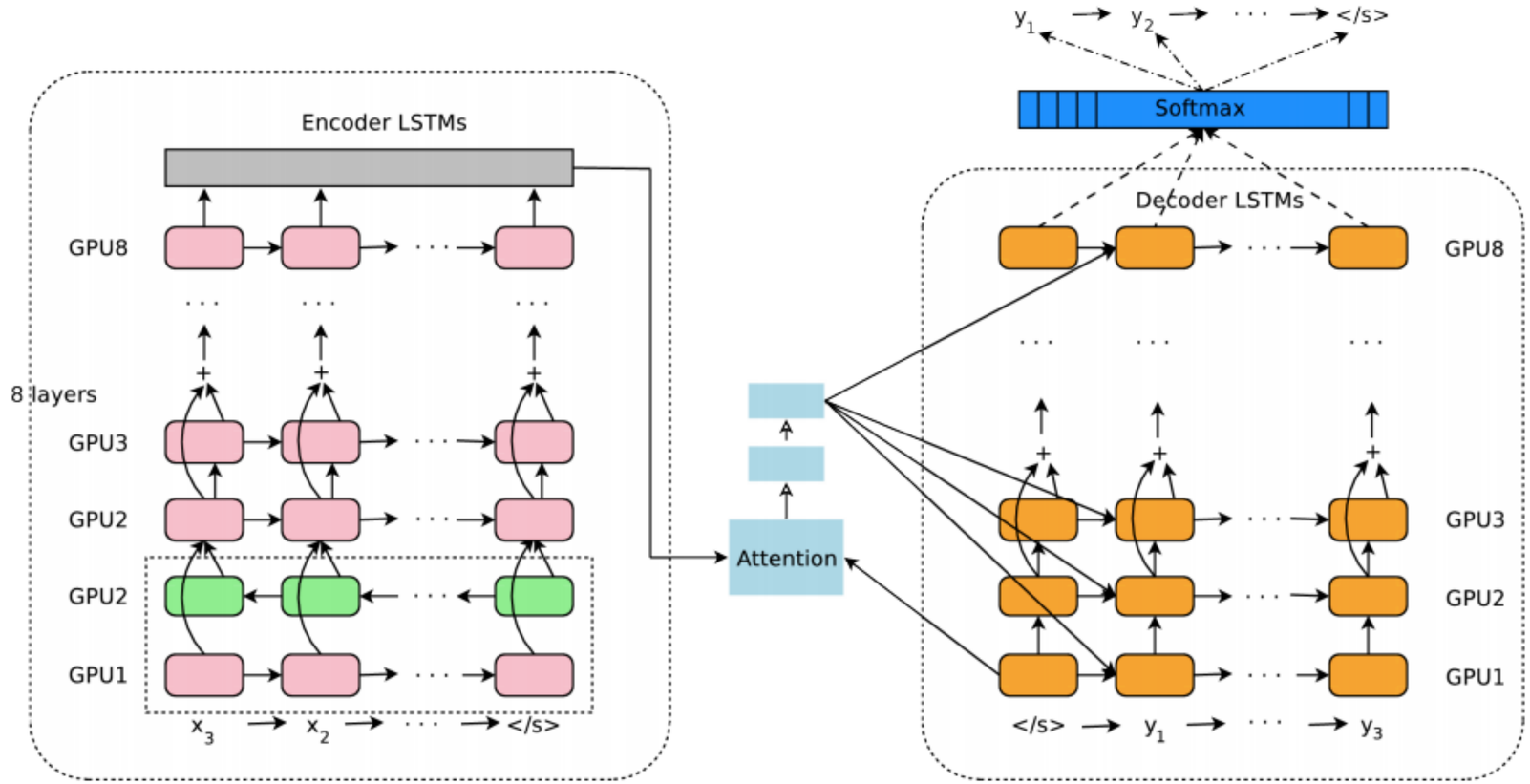
Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone
https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html



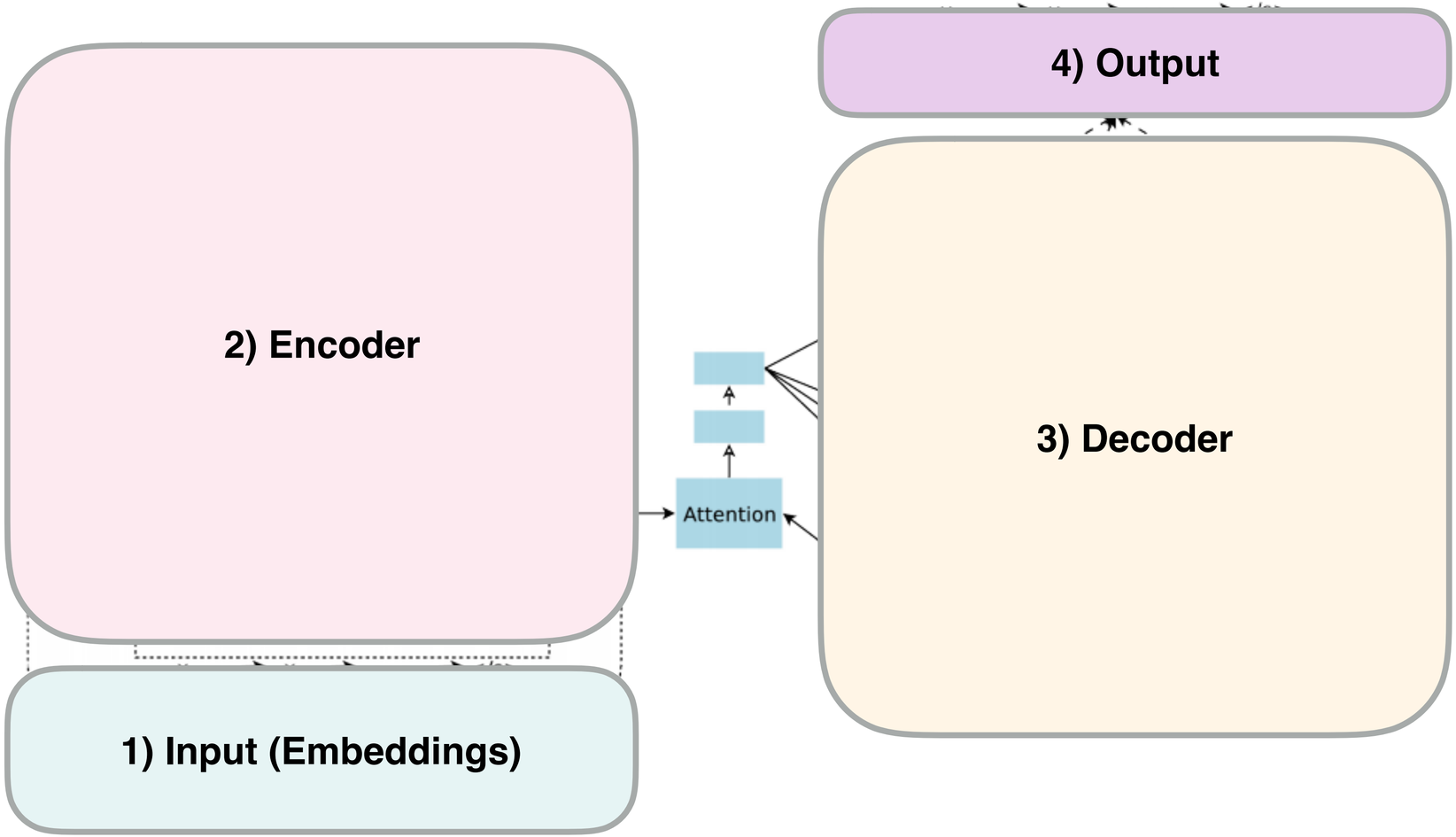
Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone
https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html



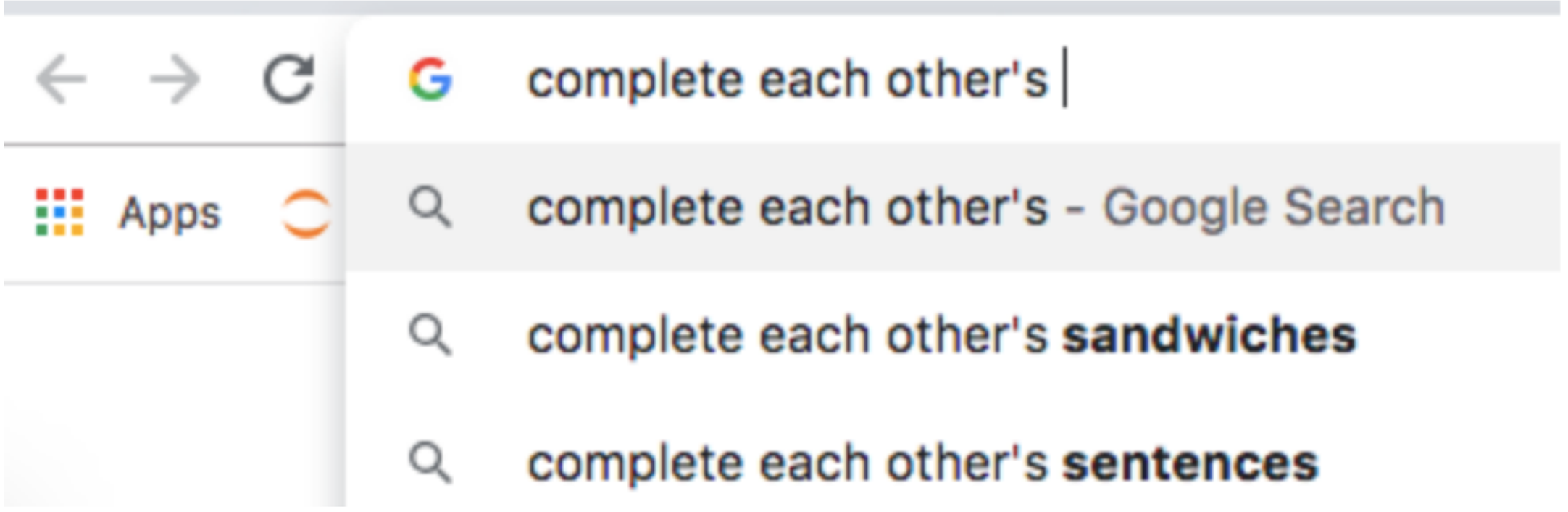
Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone
https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html



Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation (2016)
https://arxiv.org/abs/1609.08144



Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation (2016)
https://arxiv.org/abs/1609.08144


Apps
complete each other's
complete each other's - Google Search
complete each other's sandwiches
complete each other's sentences

SYSTEM PROMPT (HUMAN-WRITTEN)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them – they were so close they could touch their horns.

SYSTEM PROMPT (HUMAN-WRITTEN)

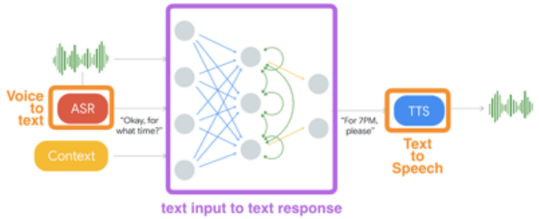
Recycling is good for the world.

NO! YOU COULD NOT BE MORE WRONG!!

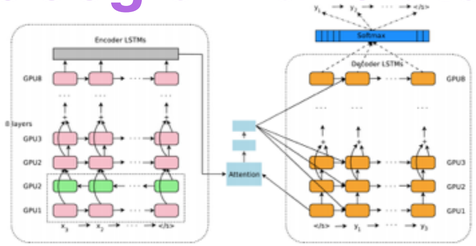
MODEL COMPLETION (MACHINE-WRITTEN, 25 TRIES)

Recycling is NOT good for the world. It is bad for the environment, it is bad for our health, and it is bad for our economy. I'm not kidding. Recycling is not good for the environment. It is destructive to the earth and it is a major contributor to global warming. Recycling is not good for our health. It contributes to obesity and diseases like heart disease and cancer. Recycling is bad for our economy. It increases the cost of a product, and in turn, the price of everything that is made with that product. Recycling is not good for our nation. We pay a tremendous price for the privilege of having the world's most advanced and efficient recycling system. Recycling is a huge, colossal waste of time, energy, money, and resources. And

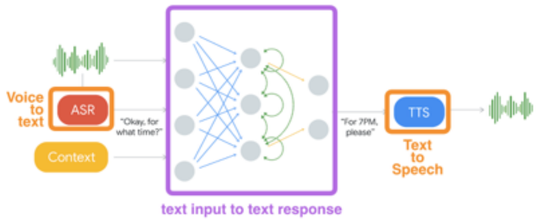
Google Duplex



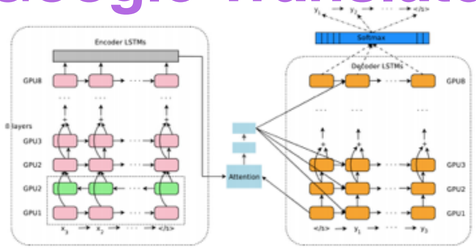
Google Translate



Google Duplex



Google Translate



THE TRANSFORMER

ULM-FiT

BERT

OpenAI Transformer

ELMo

Embeddings

SECTIONS

Intro

1. Personality Embedding
2. Word Embedding Properties

in NLP

3. Language Modeling
4. Language Modeling Training
5. Skipgram
6. Negative sampling
7. Sexist embeddings?

Non-NLP

8. Airbnb product encoding
9. Alibaba Recommendations
10. ASOS Lifetime Value
11. Other applications
12. resources

13. Consequences

Theme

Dune (1965 novel)
+ five sequels

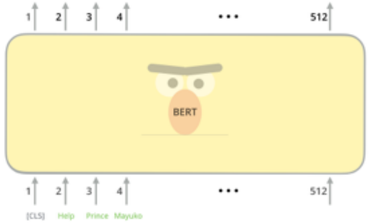
jalammar.github.io

Twitter: **jalammar**

Jay Alammar

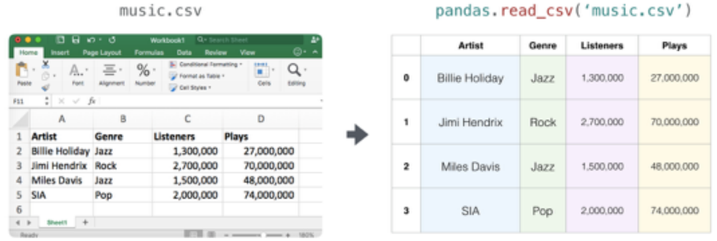
Visualizing machine learning one concept at a time

The Illustrated BERT, ELMo, and co. (How NLP Cracked Transfer Learning)



The Illustrated Transformer

A Gentle Visual Intro to Data Analysis in Python Using Pandas



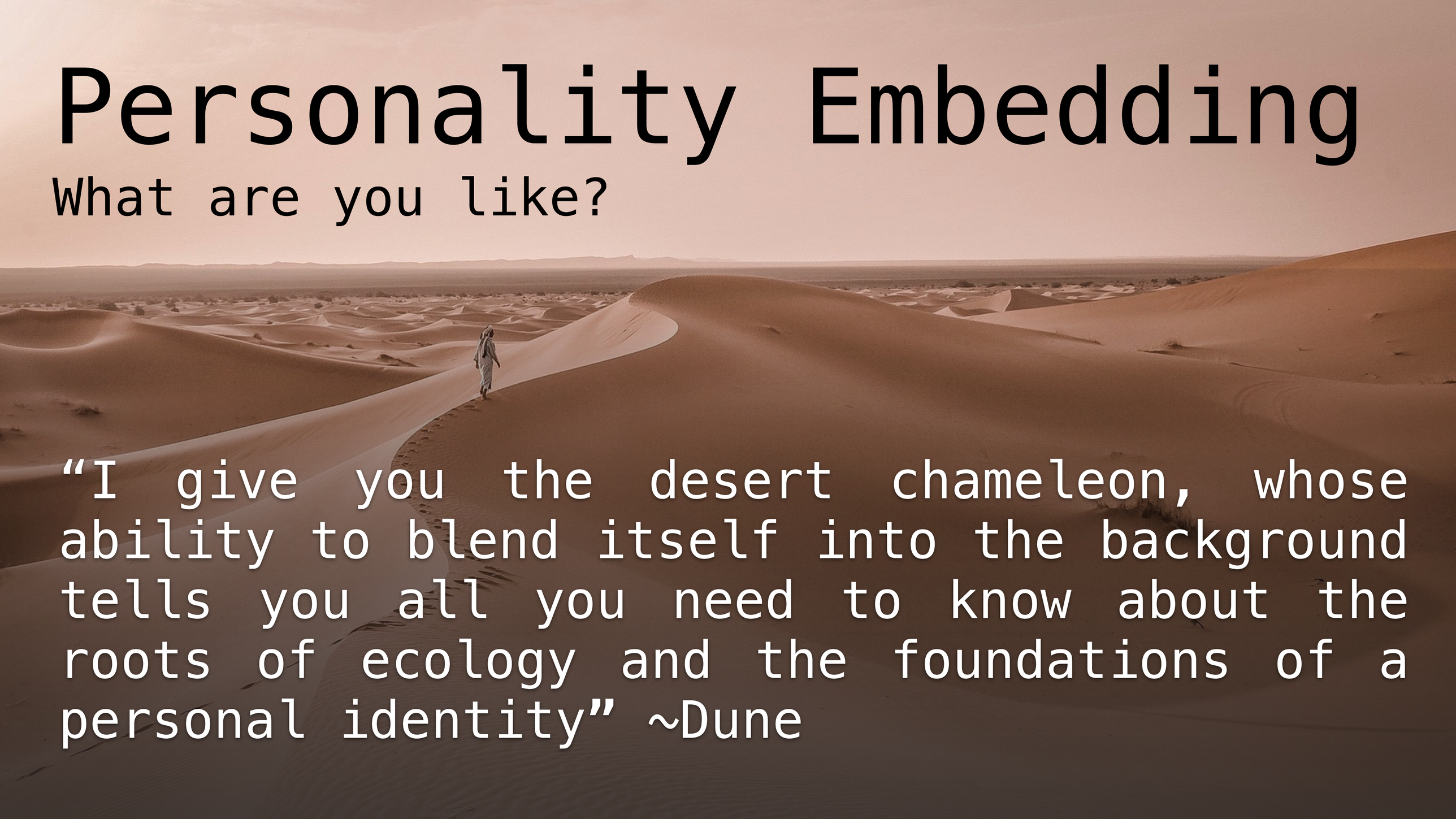
A Visual and Interactive Guide to the Basics of Neural Networks

Instructor on Udacity AI programs:

- Machine Learning Engineer Nanodegree
- Deep Learning Nanodegree
- Natural Language Processing Nanodegree

Personality Embedding

What are you like?



“I give you the desert chameleon, whose ability to blend itself into the background tells you all you need to know about the roots of ecology and the foundations of a personal identity” ~Dune

Which Kardashian Are You?

Bible, guys. YOLO

Posted on August 21, 2013, at 9:07 p.m.

Mackenzie Kruvant
BuzzFeed Staff

1. How do you feel about other people's success?

Myers-Briggs Type Indicator

Big Five Personality Traits

Openness

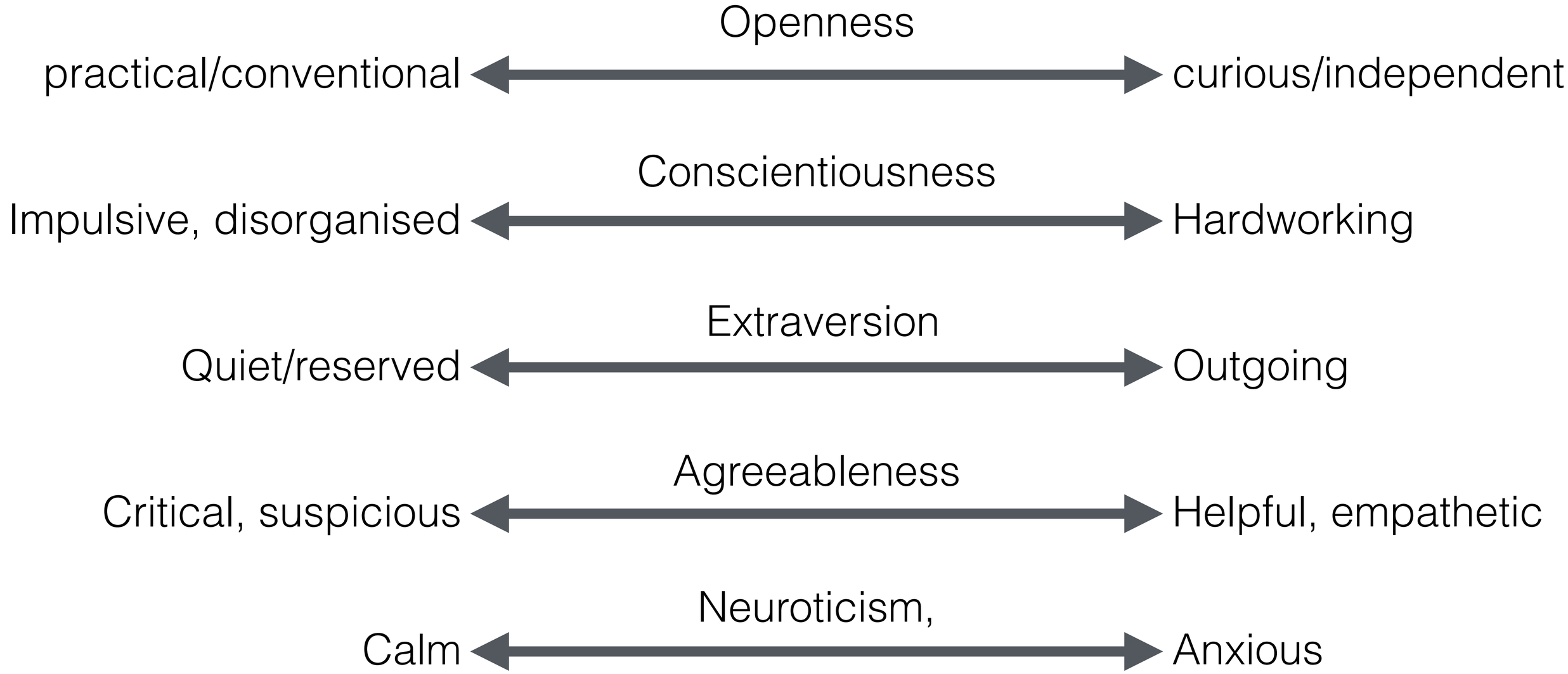
Conscientiousness

Extraversion

Agreeableness

Neuroticism,

Big Five Personality Traits



FiveThirtyEight

0 of **30** questions completed

PUBLISHED JAN. 16, 2019, AT 9:00 AM

Most Personality Quizzes Are Junk Science. Take One That Isn't.

Compare your results to those of your friends and family.

By Maggie Koerth-Baker and Julia Wolfe

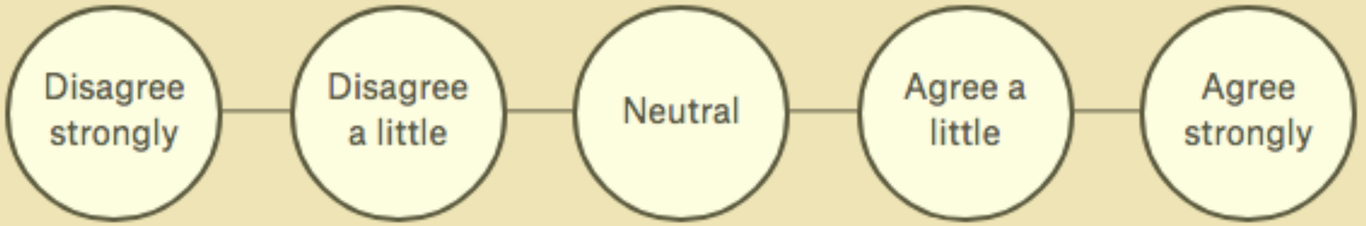
What's your personality, and what can it tell you about your true self? Those questions have launched a thousand online personality quizzes. But you can

https://projects.fivethirtyeight.com/personality-quiz/

1 of **30** questions completed

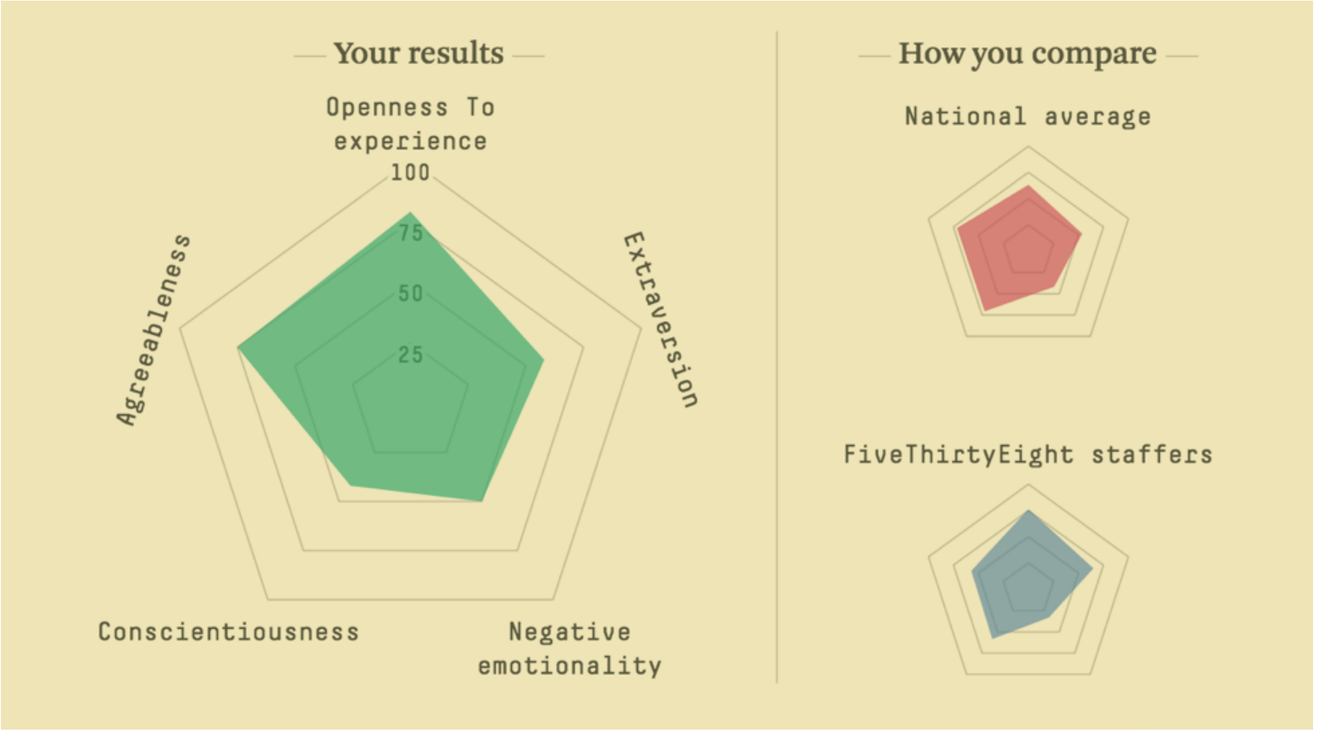
I tend to be quiet.

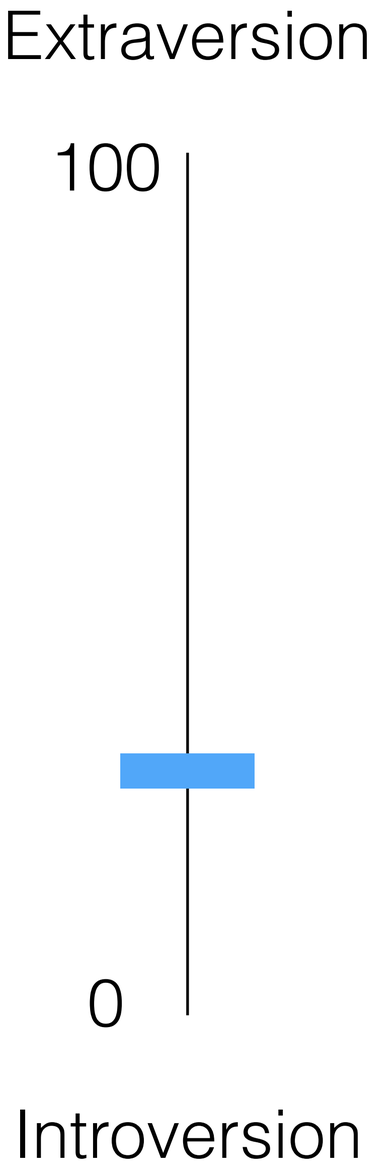
I am compassionate and have a soft heart.



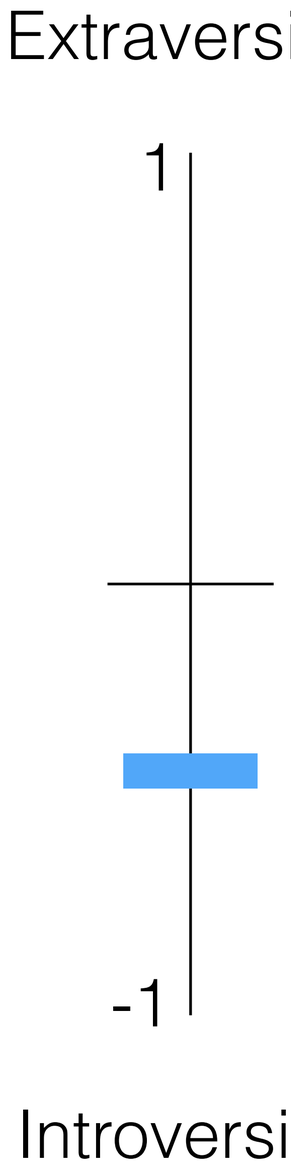
Openness to experience	79 out of 100
Agreeableness	75 out of 100
Conscientiousness	42 out of 100
Negative emotionality	50 out of 100
Extraversion	58 out of 100

Openness to experience	79 out of 100
Agreeableness	75 out of 100
Conscientiousness	42 out of 100
Negative emotionality	50 out of 100
Extraversion	58 out of 100

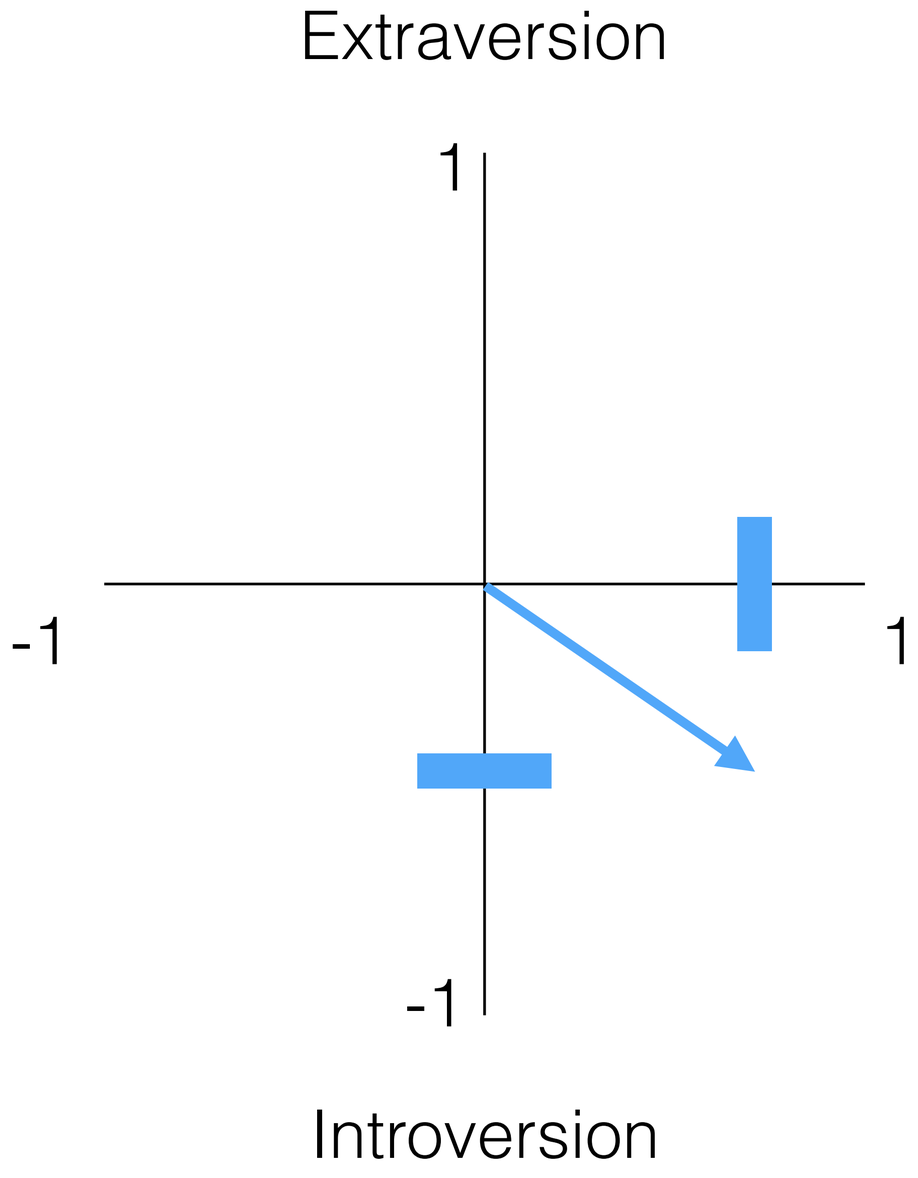




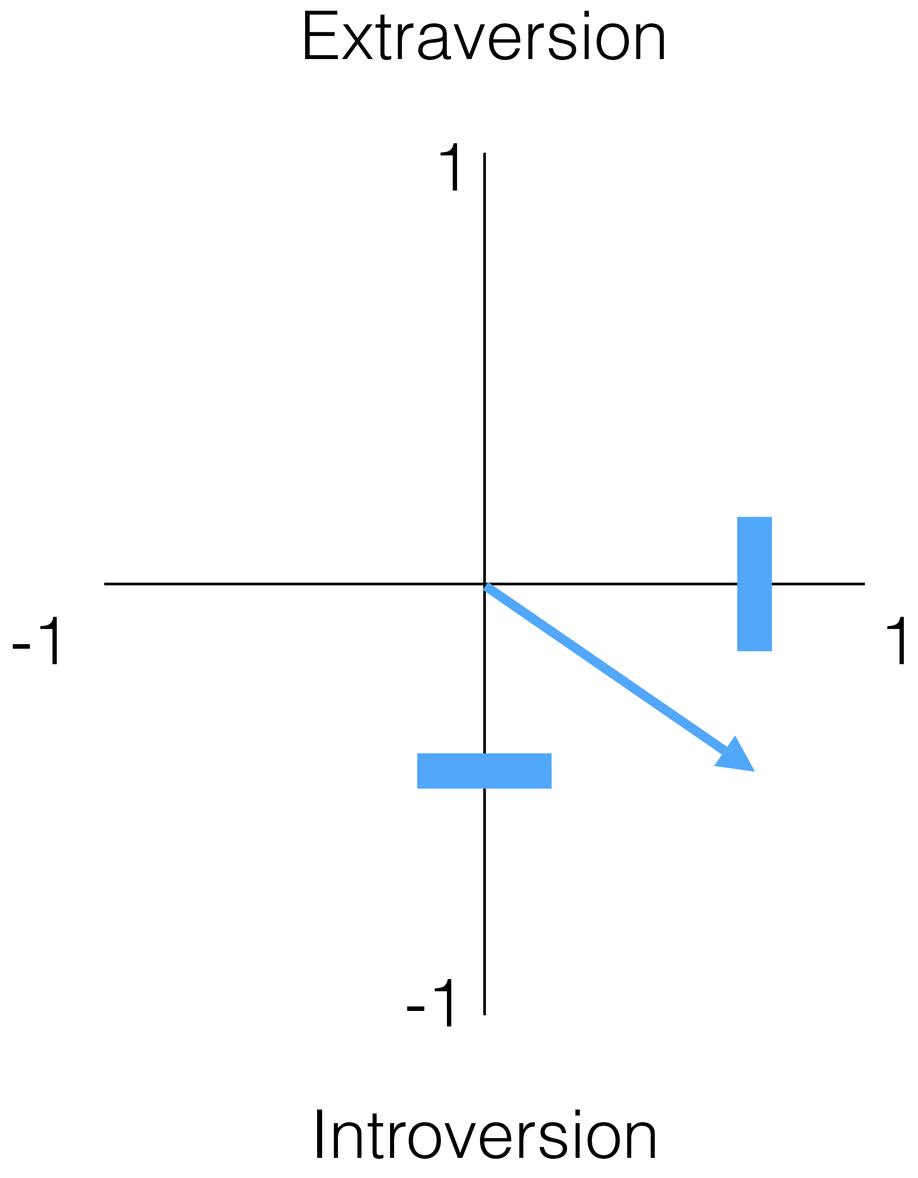
	Extraversion
Jay	38



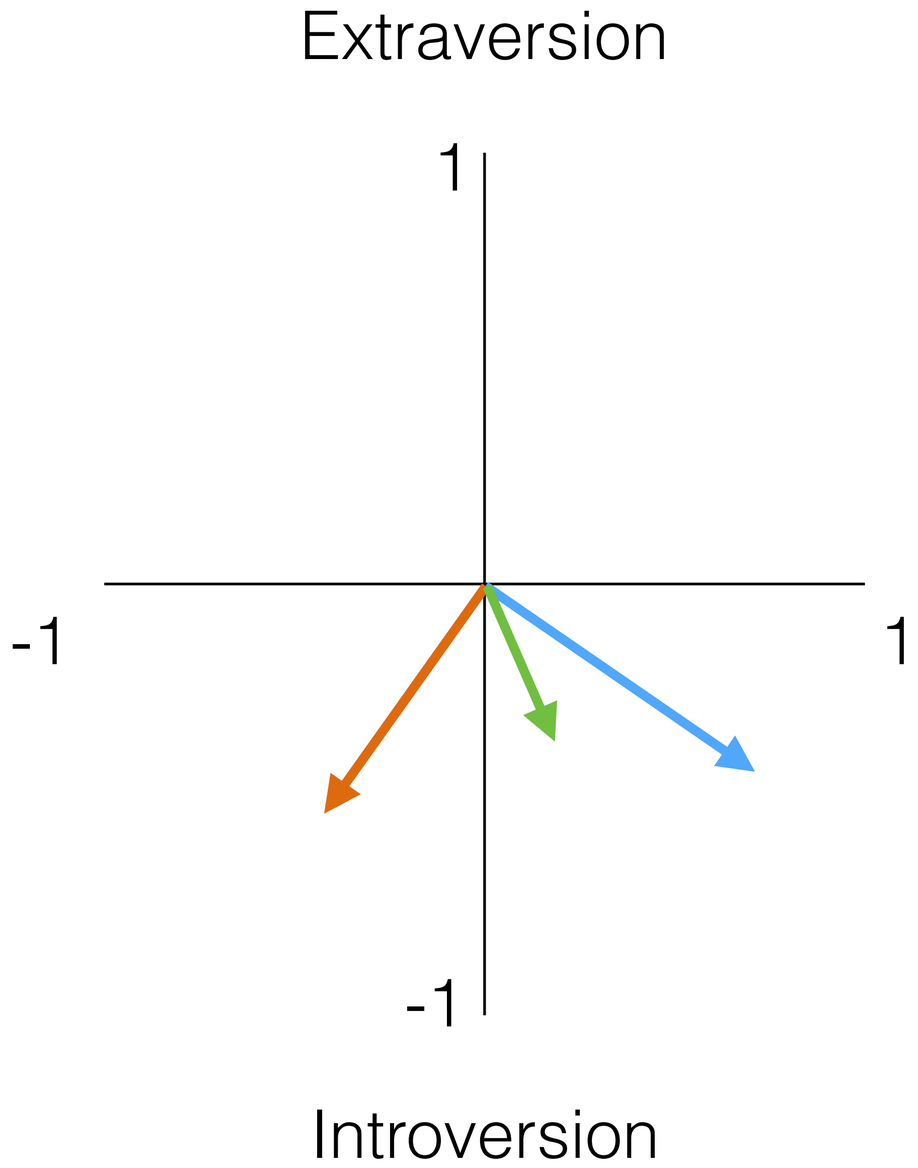
	Extraversion
Jay	-0.4



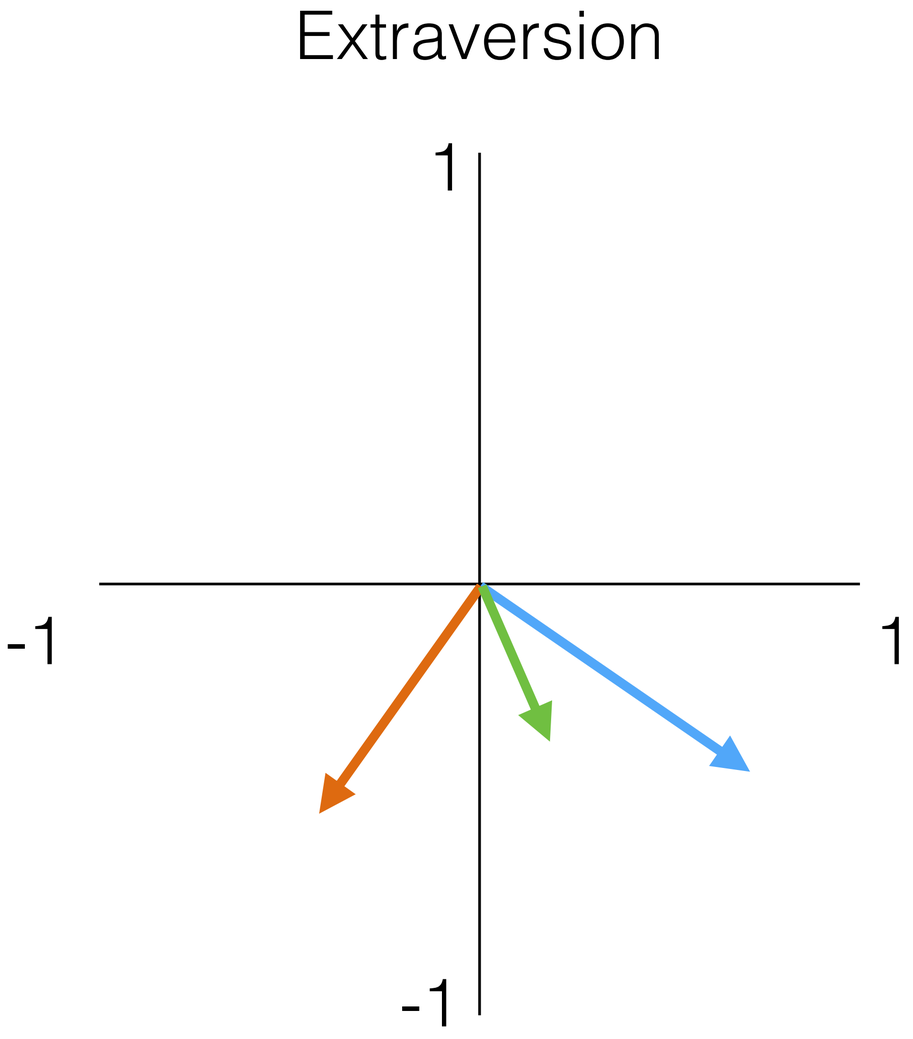
	Trait #1	Trait #2			
Jay	-0.4	0.8			



	Trait #1	Trait #2			
Jay	-0.4	0.8			



Jay	-0.4				
Person #1	-0.3	0.2			
Person #2	-0.5	-0.4			



Jay

-0.4				

Person #1

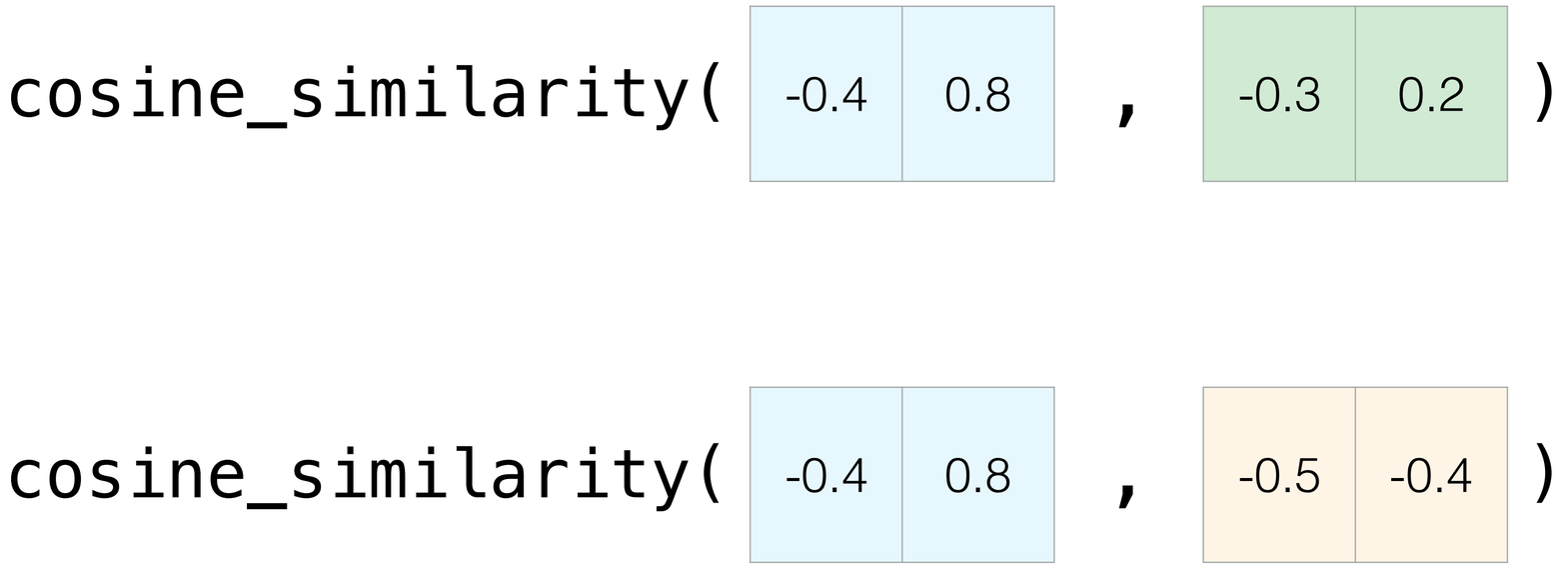
-0.3	0.2			

Person #2

-0.5	-0.4			

Similarity(| | | | | | , | | | | | |)

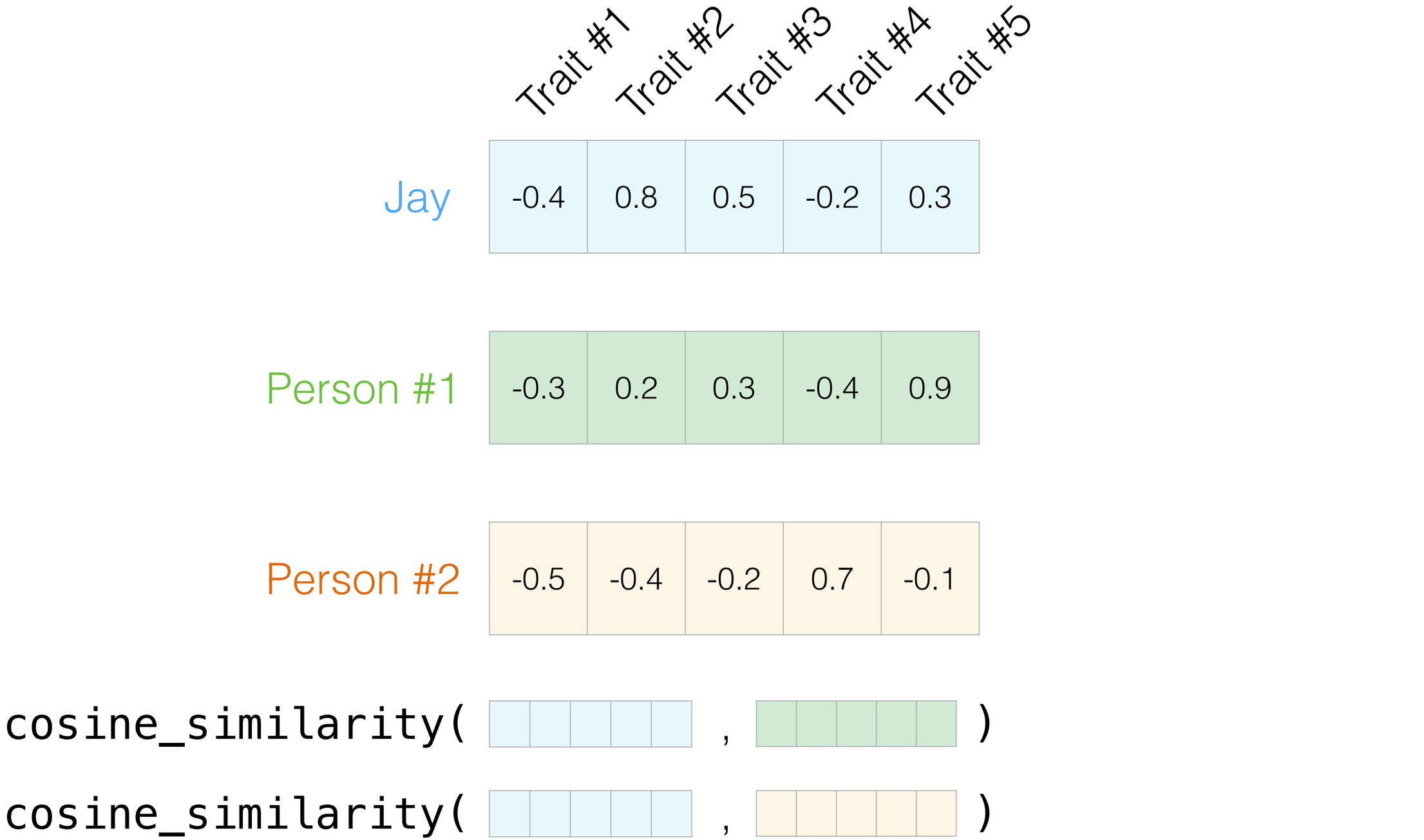
Similarity(| | | | | | , | | | | | |)


cosine_similarity(-0.4 0.8 , -0.3 0.2)
cosine_similarity(-0.4 0.8 , -0.5 -0.4)

cosine_similarity(| -0.4 | 0.8 | , | -0.3 | 0.2 |) = 0.87

cosine_similarity(| -0.4 | 0.8 | , | -0.5 | -0.4 |) = −0.20

	Trait #1	Trait #2	Trait #3	Trait #4	Trait #5
Jay	-0.4	0.8	0.5	-0.2	0.3
Person #1	-0.3	0.2	0.3	-0.4	0.9
Person #2	-0.5	-0.4	-0.2	0.7	-0.1


Trait #1
Trait #2
Trait #3
Trait #4
Trait #5
Jay
-0.4
0.8
0.5
-0.2
0.3
Person #1
-0.3
0.2
0.3
-0.4
0.9
Person #2
-0.5
-0.4
-0.2
0.7
-0.1
cosine_similarity(
,
)
cosine_similarity(
,
)

1- We can represent things (and people) as vectors of numbers
(Which is great for machines!)

Jay	-0.4	0.8	0.5	-0.2	0.3

1- We can represent things (and people) as vectors of numbers
(Which is great for machines!)

Jay	-0.4	0.8	0.5	-0.2	0.3

2- We can easily calculate how similar vectors are to each other

1- We can represent things (and people) as vectors of numbers
(Which is great for machines!)

Jay	-0.4	0.8	0.5	-0.2	0.3

2- We can easily calculate how similar vectors are to each other

Customers who liked Jay also liked:

	cosine_similarity ▼
Person #1	0.86
Person #2	0.5
Person #3	-0.20

Word Embeddings

“The gift of words is the gift of deception and illusion”
~ Dune

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

John Rupert Firth

“king”

```
array([ 0.50451 ,  0.68607 , -0.59517 , -0.022801,  0.60046 , -0.13498 ,
       -0.08813 ,  0.47377 , -0.61798 , -0.31012 , -0.076666,  1.493   ,
       -0.034189, -0.98173 ,  0.68229 ,  0.81722 , -0.51874 , -0.31503 ,
       -0.55809 ,  0.66421 ,  0.1961  , -0.13495 , -0.11476 , -0.30344 ,
        0.41177 , -2.223   , -1.0756  , -1.0783  , -0.34354 ,  0.33505 ,
        1.9927  , -0.04234 , -0.64319 ,  0.71125 ,  0.49159 ,  0.16754 ,
        0.34344 , -0.25663 , -0.8523  ,  0.1661  ,  0.40102 ,  1.1685  ,
       -1.0137  , -0.21585 , -0.15155 ,  0.78321 , -0.91241 , -1.6106  ,
       -0.64426 , -0.51042 ], dtype=float32)
```

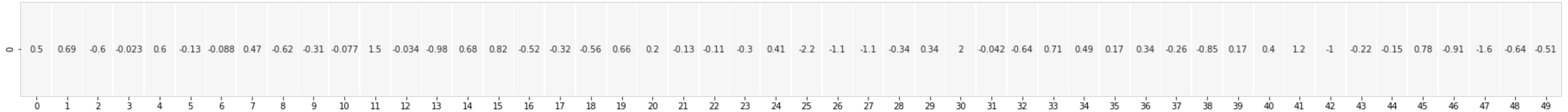
“king”

```
array([ 0.50451 ,  0.68607 , -0.59517 , -0.022801,  0.60046 , -0.13498 ,
       -0.08813 ,  0.47377 , -0.61798 , -0.31012 , -0.076666,  1.493   ,
       -0.034189, -0.98173 ,  0.68229 ,  0.81722 , -0.51874 , -0.31503 ,
       -0.55809 ,  0.66421 ,  0.1961  , -0.13495 , -0.11476 , -0.30344 ,
        0.41177 , -2.223   , -1.0756  , -1.0783  , -0.34354 ,  0.33505 ,
        1.9927  , -0.04234 , -0.64319 ,  0.71125 ,  0.49159 ,  0.16754 ,
        0.34344 , -0.25663 , -0.8523  ,  0.1661  ,  0.40102 ,  1.1685  ,
       -1.0137  , -0.21585 , -0.15155 ,  0.78321 , -0.91241 , -1.6106  ,
       -0.64426 , -0.51042 ], dtype=float32)
```

Embedding: GloVe
Dimensions: 50
Trained on: Wikipedia + Gigaword 5 (6B tokens)
Number of vectors: 400,000

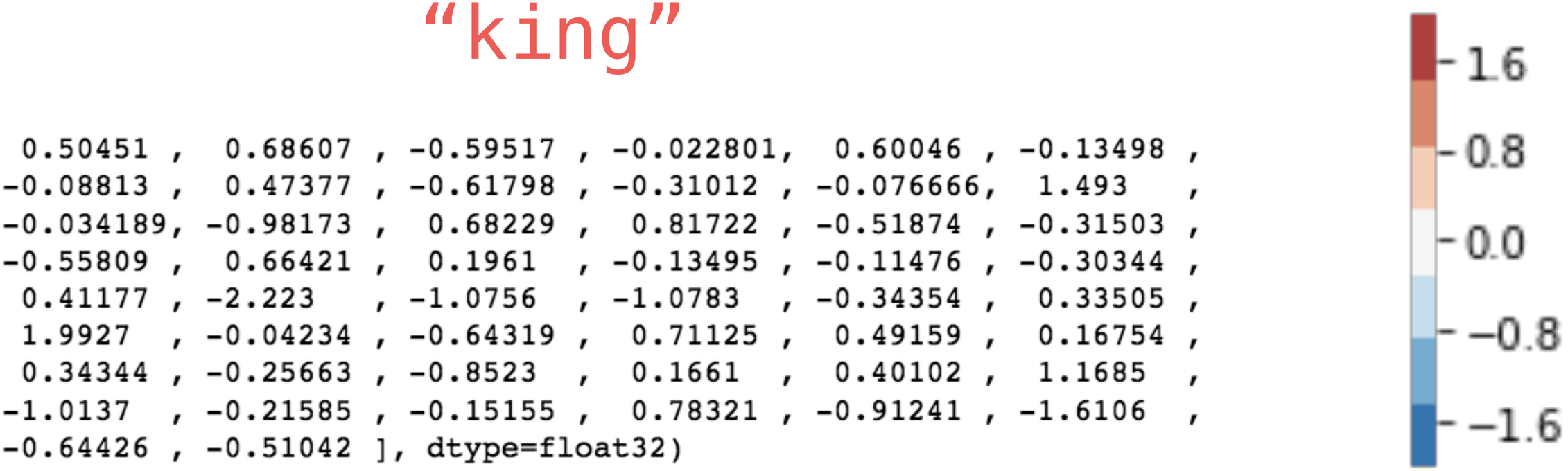
“king”

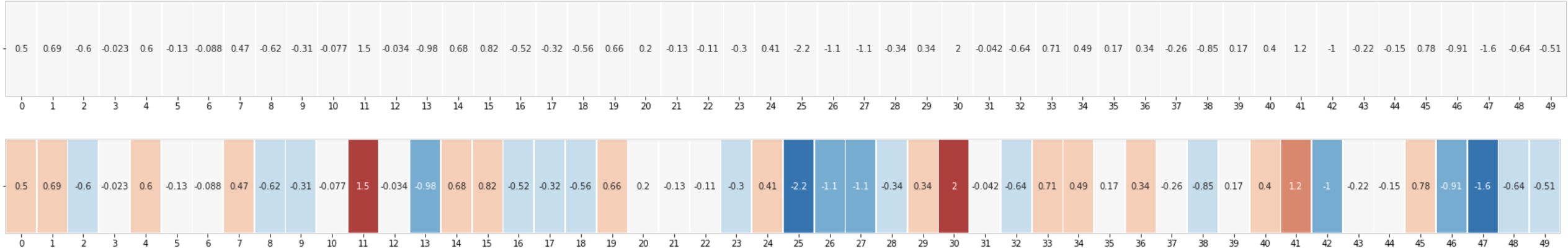
```
array([ 0.50451 ,  0.68607 , -0.59517 , -0.022801,  0.60046 , -0.13498 ,
       -0.08813 ,  0.47377 , -0.61798 , -0.31012 , -0.076666,  1.493   ,
       -0.034189, -0.98173 ,  0.68229 ,  0.81722 , -0.51874 , -0.31503 ,
       -0.55809 ,  0.66421 ,  0.1961  , -0.13495 , -0.11476 , -0.30344 ,
        0.41177 , -2.223   , -1.0756  , -1.0783  , -0.34354 ,  0.33505 ,
        1.9927  , -0.04234 , -0.64319 ,  0.71125 ,  0.49159 ,  0.16754 ,
        0.34344 , -0.25663 , -0.8523  ,  0.1661  ,  0.40102 ,  1.1685  ,
       -1.0137  , -0.21585 , -0.15155 ,  0.78321 , -0.91241 , -1.6106  ,
       -0.64426 , -0.51042 ], dtype=float32)
```



“king”

```
array([ 0.50451 ,  0.68607 , -0.59517 , -0.022801,  0.60046 , -0.13498 ,
       -0.08813 ,  0.47377 , -0.61798 , -0.31012 , -0.076666,  1.493   ,
       -0.034189, -0.98173 ,  0.68229 ,  0.81722 , -0.51874 , -0.31503 ,
       -0.55809 ,  0.66421 ,  0.1961  , -0.13495 , -0.11476 , -0.30344 ,
        0.41177 , -2.223   , -1.0756  , -1.0783  , -0.34354 ,  0.33505 ,
        1.9927  , -0.04234 , -0.64319 ,  0.71125 ,  0.49159 ,  0.16754 ,
        0.34344 , -0.25663 , -0.8523  ,  0.1661  ,  0.40102 ,  1.1685  ,
       -1.0137  , -0.21585 , -0.15155 ,  0.78321 , -0.91241 , -1.6106  ,
       -0.64426 , -0.51042 ], dtype=float32)
```

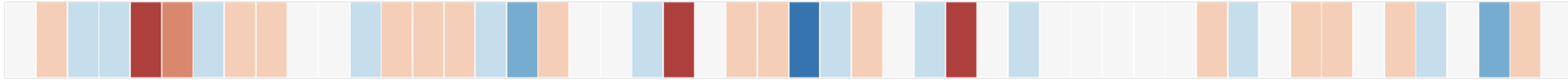


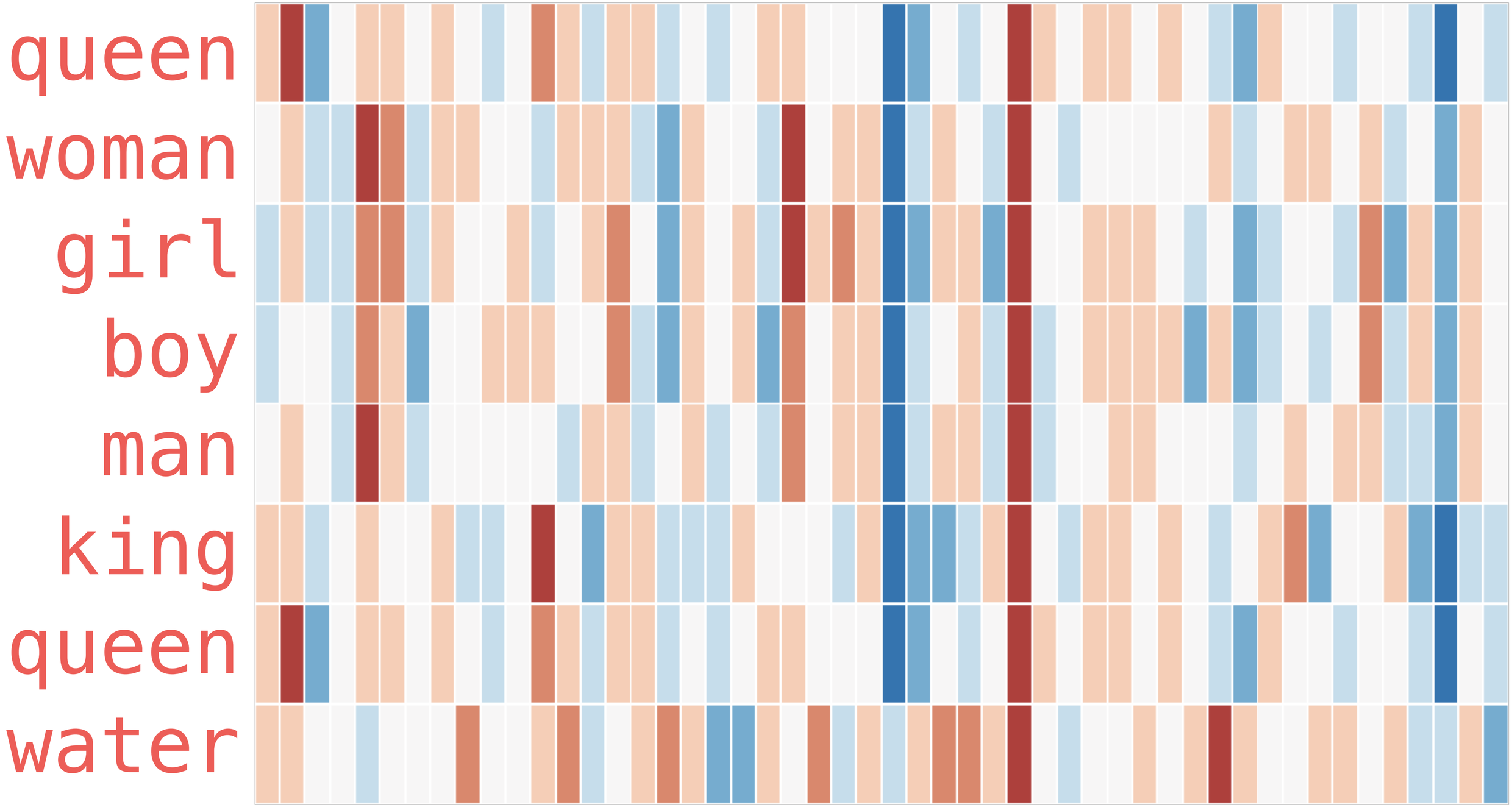


“king”

“Man”

“Woman”




queen
woman
girl
boy
man
king
queen
water

Analogies

king − man + woman ~=

king - man + woman ~= queen

```
model.most_similar(positive=["king","woman"], negative=["man"])
```

```
[('queen', 0.8523603677749634),
 ('throne', 0.7664333581924438),
 ('prince', 0.7592144012451172),
 ('daughter', 0.7473883032798767),
 ('elizabeth', 0.7460219860076904),
 ('princess', 0.7424570322036743),
 ('kingdom', 0.7337411642074585),
 ('monarch', 0.721449077129364),
 ('eldest', 0.7184862494468689),
 ('widow', 0.7099430561065674)]
```

king − man + woman ~= queen

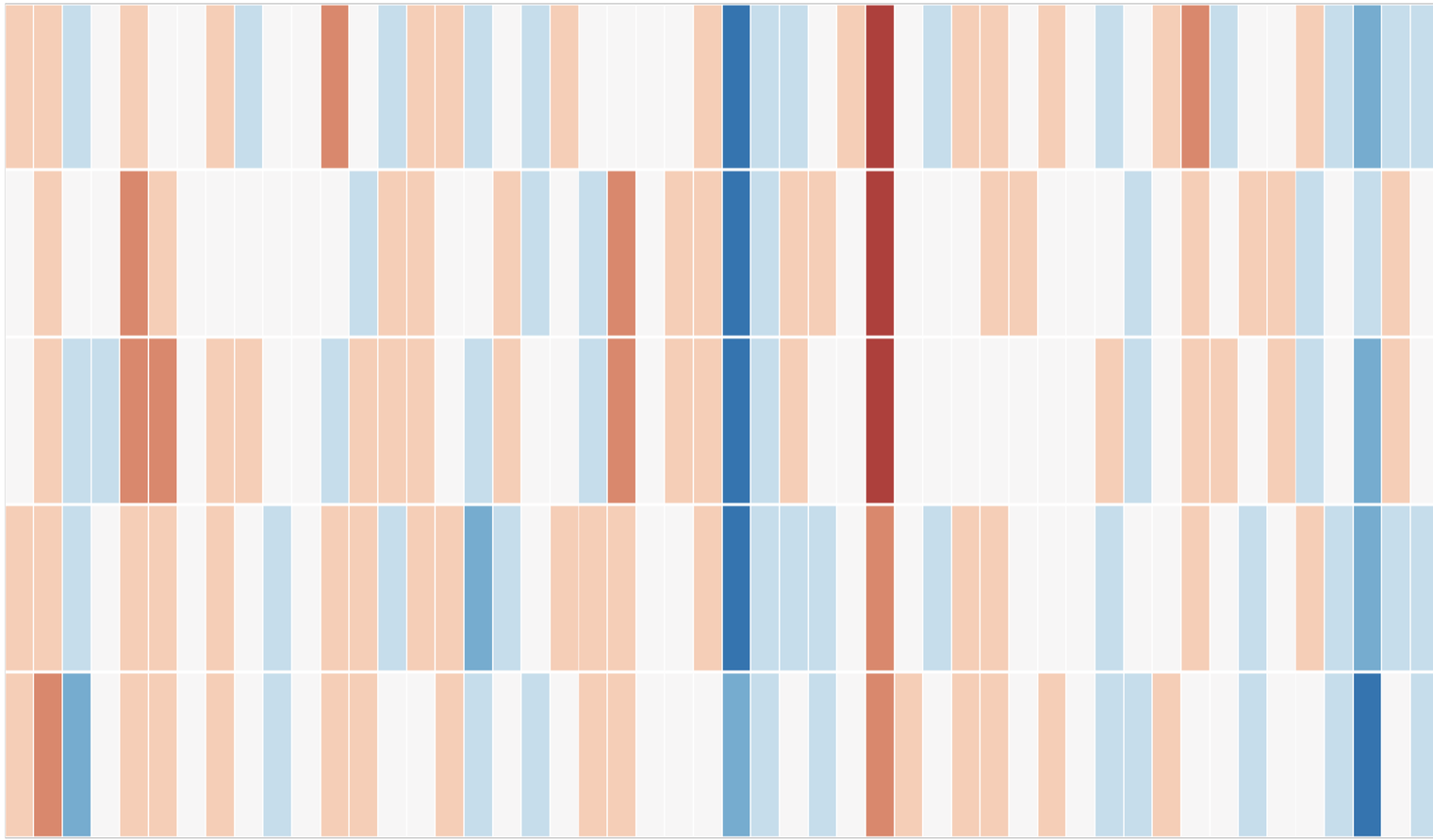
king

man

woman

king−man+woman

queen

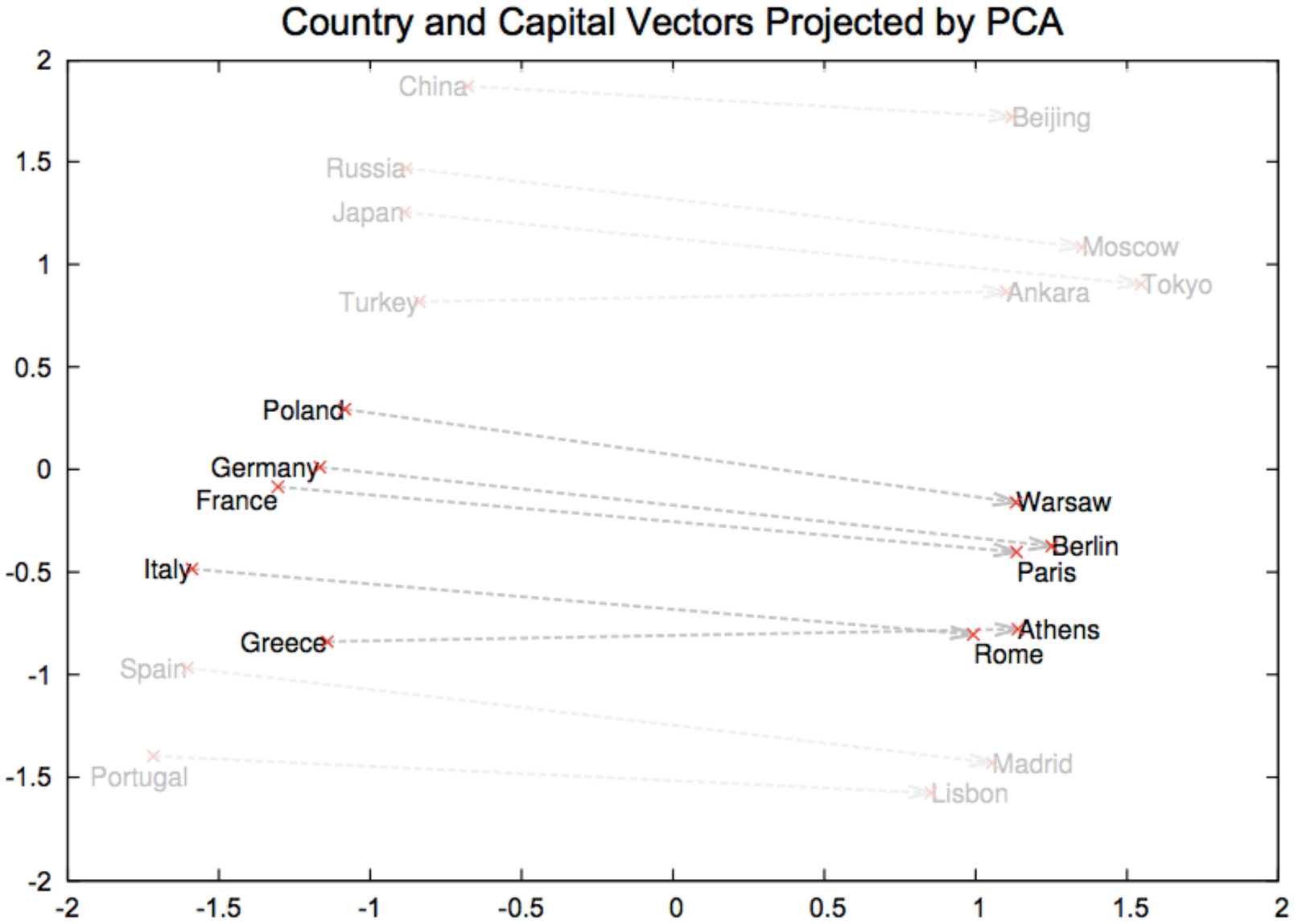


france is to paris as italy is to _____

```
model.most_similar(positive=["paris", "italy"], negative=["france"])
```

```
[('rome', 0.8465589284896851),
 ('milan', 0.77660071849823),
 ('turin', 0.7666354775428772),
 ('venice', 0.7591592073440552),
 ('madrid', 0.756580471992492 7),
 ('italian', 0.7513630390167236),
 ('aires', 0.7429094910621643),
 ('naples', 0.7406353950500488),
 ('buenos', 0.7357348203659058),
 ('lisbon', 0.7244545221328735)]
```

france is to paris as italy is to _____



Distributed Representations of Words and Phrases and their Compositionality

Tomas Mikolov
Google Inc.
Mountain View
mikolov@google.com

Ilya Sutskever
Google Inc.
Mountain View
ilyasu@google.com

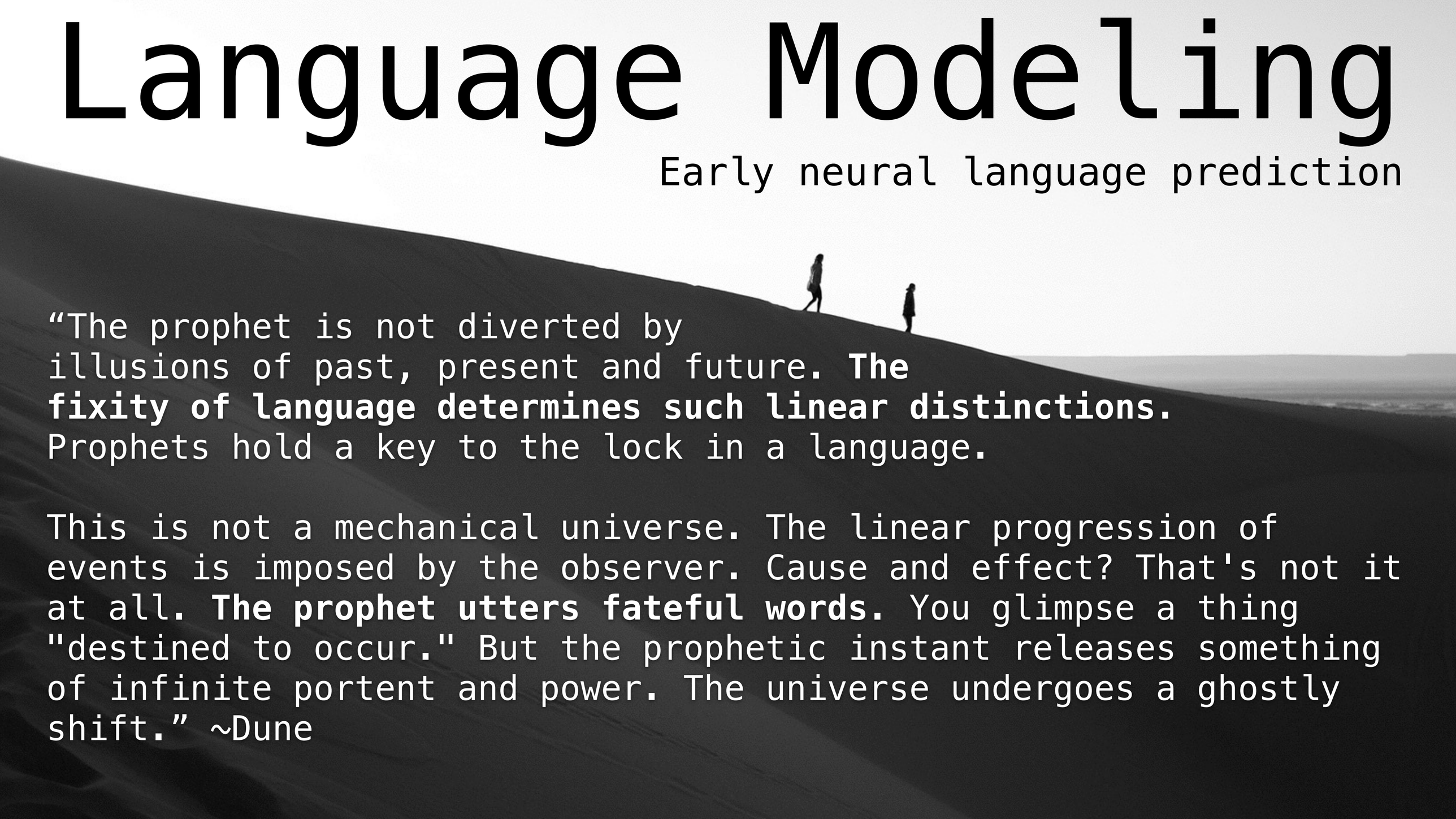
Kai Chen
Google Inc.
Mountain View
kai@google.com

Greg Corrado
Google Inc.
Mountain View
gcorrado@google.com

Jeffrey Dean
Google Inc.
Mountain View
jeff@google.com

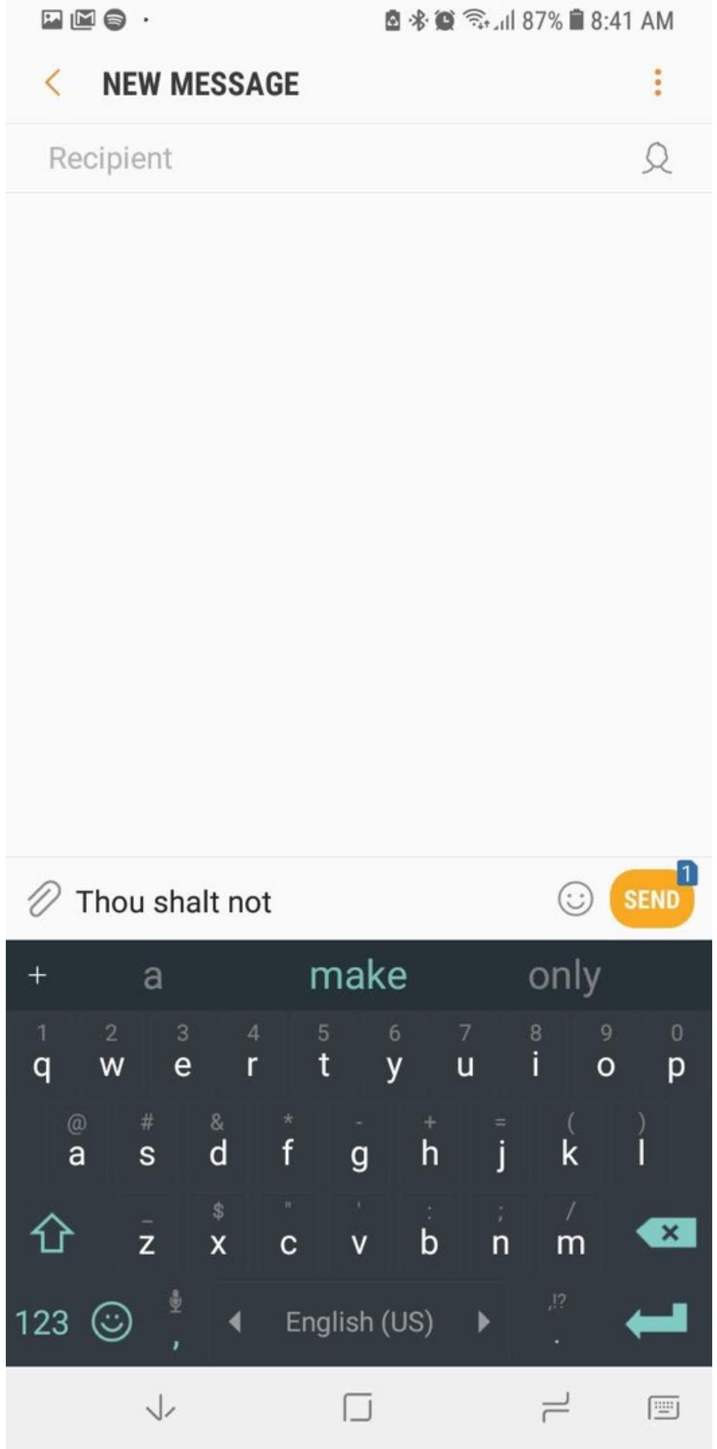
Language Modeling

Early neural language prediction

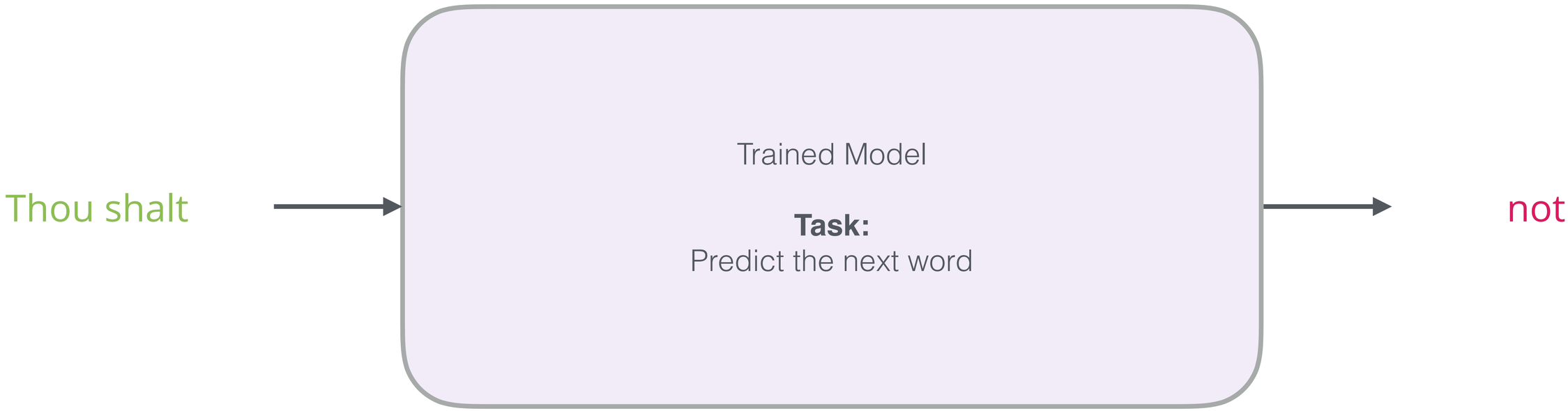


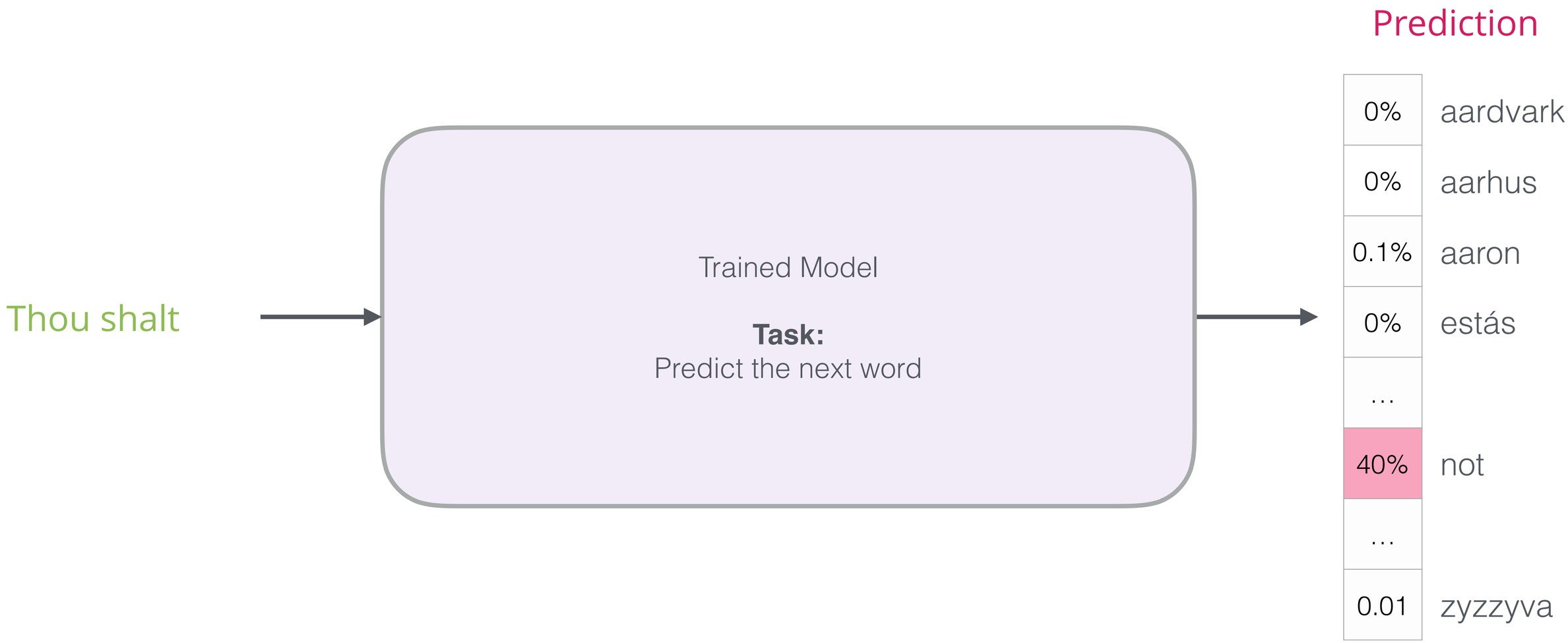
“The prophet is not diverted by illusions of past, present and future. **The fixity of language determines such linear distinctions.** Prophets hold a key to the lock in a language.

This is not a mechanical universe. The linear progression of events is imposed by the observer. Cause and effect? That's not it at all. **The prophet utters fateful words.** You glimpse a thing "destined to occur." But the prophetic instant releases something of infinite portent and power. The universe undergoes a ghostly shift.” ~Dune


87% 8:41 AM
NEW MESSAGE
Recipient
Thou shalt not
SEND
1
a
make
only
English (US)

Thou shalt _____


Thou shalt
Trained Model
Task:
Predict the next word
not


Prediction
Thou shalt
Trained Model
Task:
Predict the next word
0%
aardvark
0%
aarhus
0.1%
aaron
0%
estás
…
40%
not
…
0.01
zyzzyva

Input

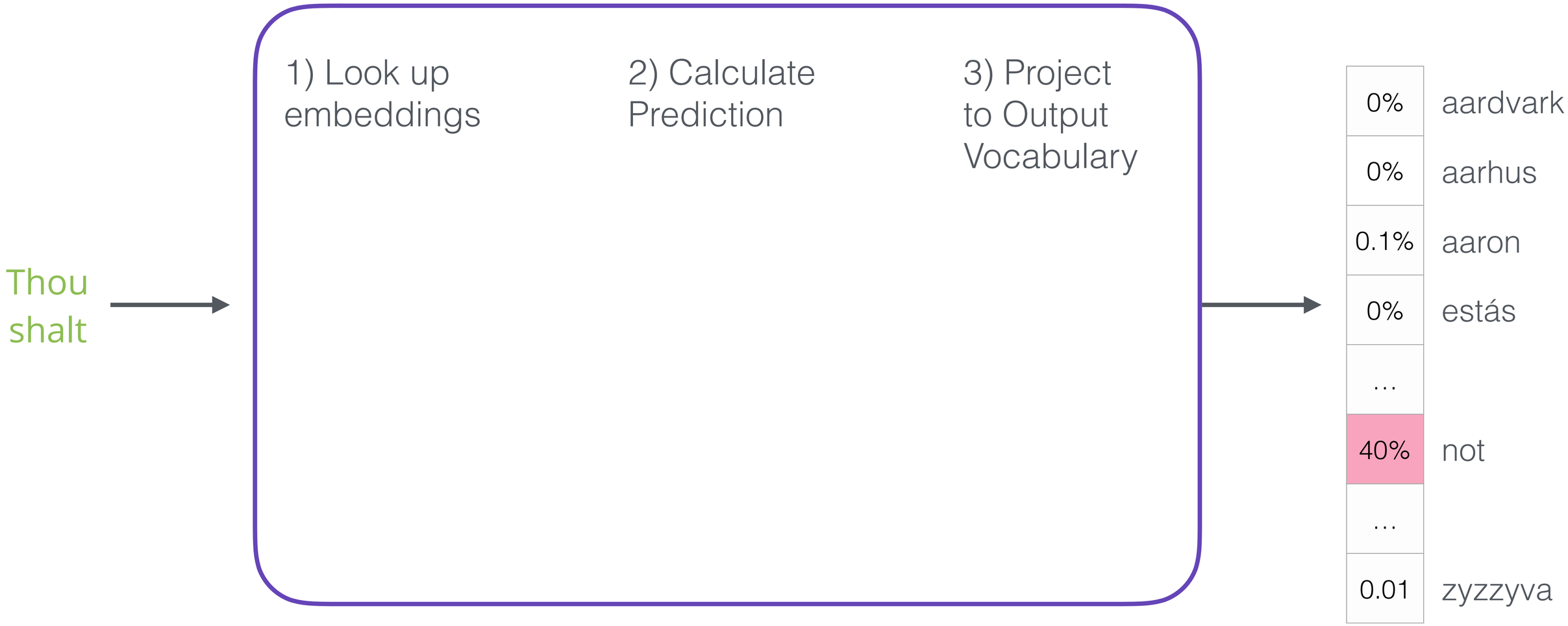
Features

Trained Model

Task:
Predict the next word

Output

Prediction



Input

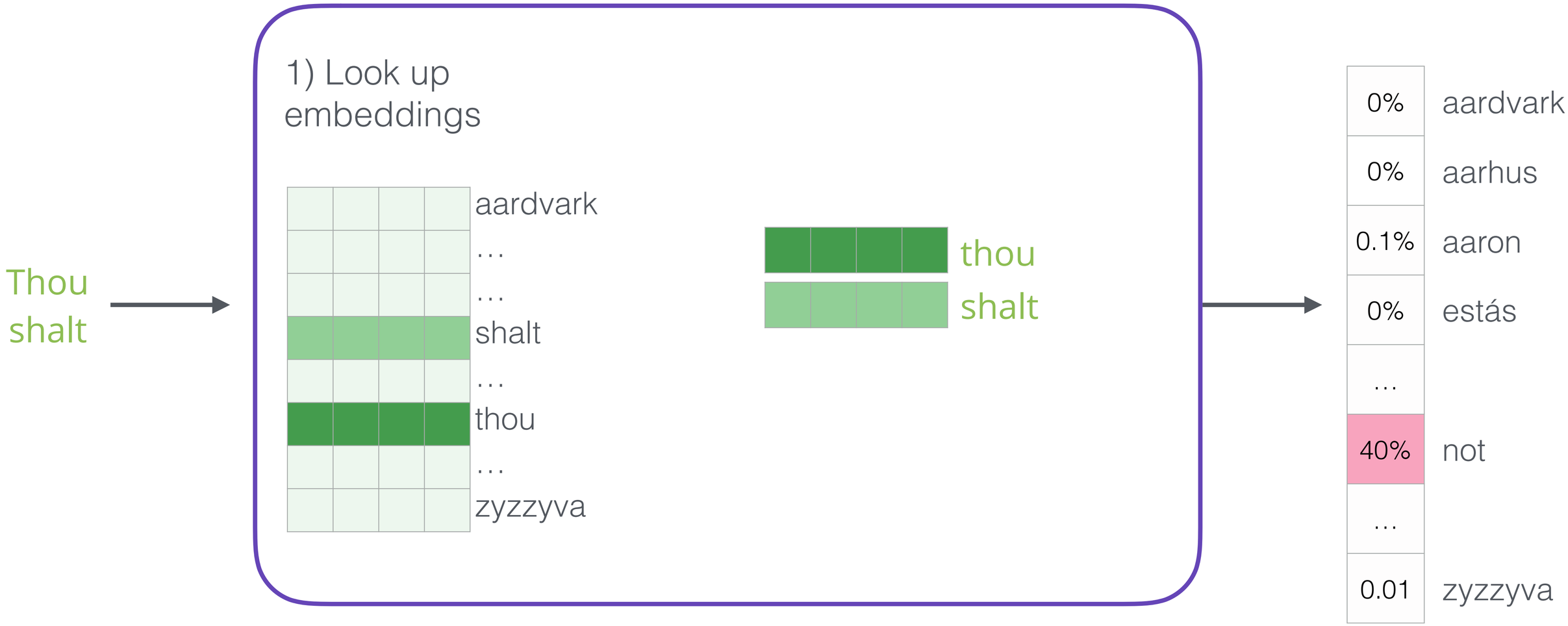
Features

Trained Model

Task:
Predict the next word

Output

Prediction



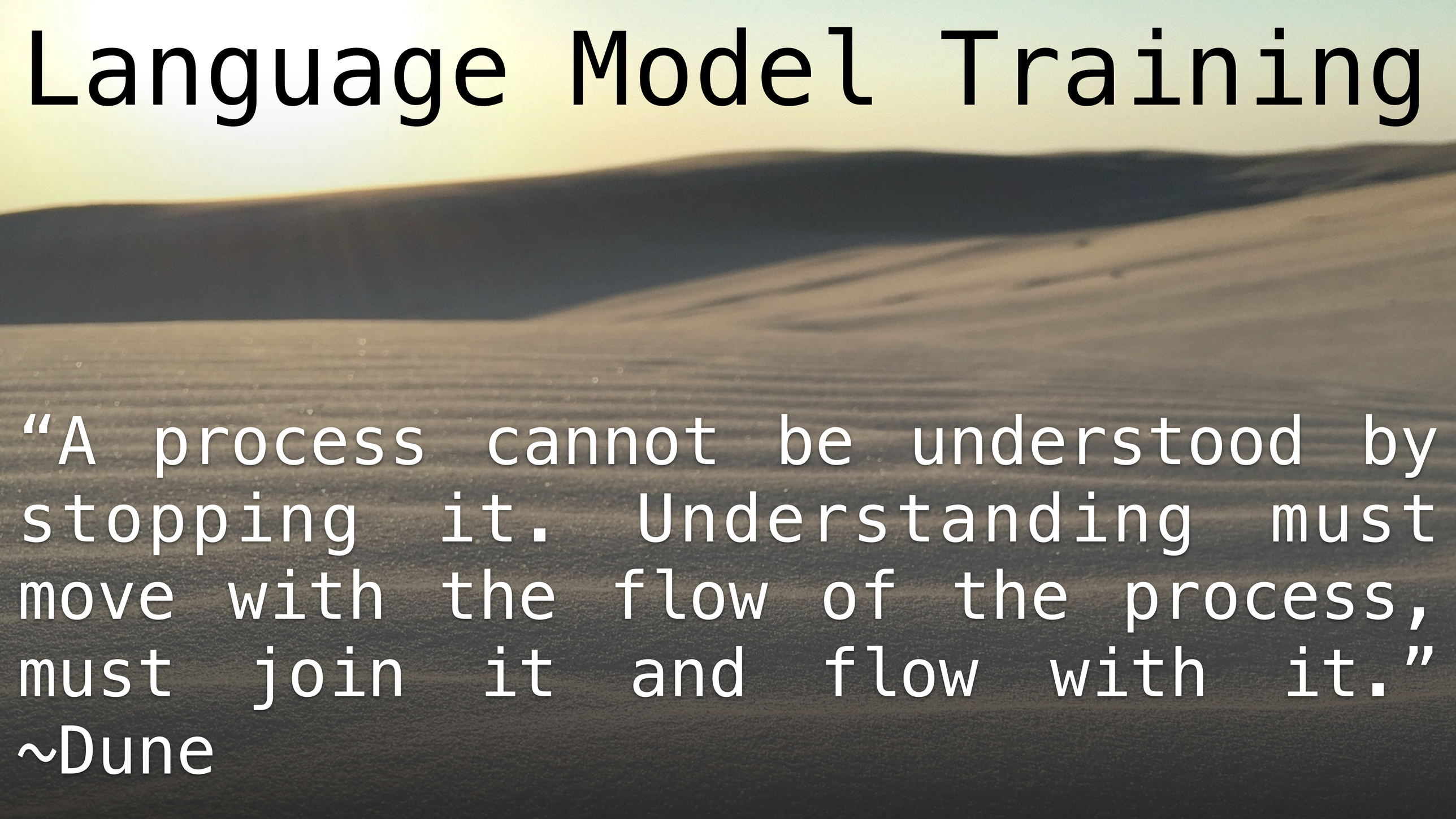
A Neural Probabilistic Language Model

Yoshua Bengio BENGIOY@IRO.UMONTREAL.CA
Réjean Ducharme DUCHARME@IRO.UMONTREAL.CA
Pascal Vincent VINCENTP@IRO.UMONTREAL.CA
Christian Jauvin JAUVINC@IRO.UMONTREAL.CA
Département d'Informatique et Recherche Opérationnelle
Centre de Recherche Mathématiques
Université de Montréal, Montréal, Québec, Canada

Editors: Jaz Kandola, Thomas Hofmann, Tomaso Poggio and John Shawe-Taylor

http://www.jmlr.org/papers/volume3/bengio03a/bengio03a.pdf

Language Model Training



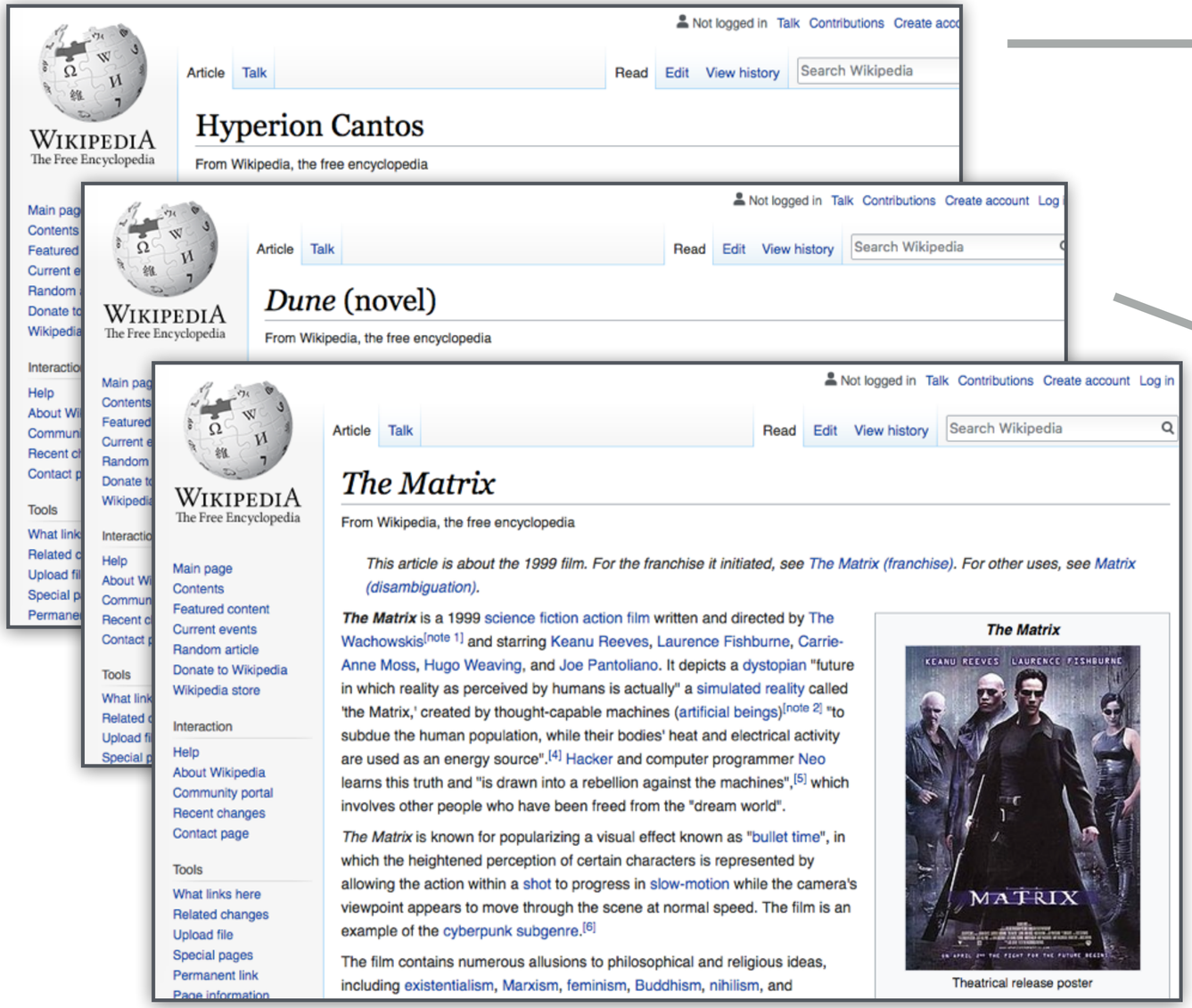
“A process cannot be understood by stopping it. Understanding must move with the flow of the process, must join it and flow with it.”
~Dune

Model: Untrained Model

Dataset:

Objective: Predict the next word

Input 1	Input 2	Output



The **Hyperion Cantos** is a series of science fiction novels by Dan Simmons. The title was originally used for the collection of the first pair of books in the series, *Hyperion* and *The Fall of Hyperion*,[1][2] and later came to refer to the overall storyline, including *Endymion*, *The Rise of Endymion*, and a number of short stories.[3][4] More narrowly, inside the fictional storyline, after the first volume, the Hyperion Cantos is an epic poem written by the character Martin Silenus covering in verse form the events of the first book.[5]

Of the four novels, *Hyperion* received the Hugo and Locus Awards in 1990;[6] *The Fall of Hyperion* won the Locus and British Science Fiction Association Awards in 1991;[7] and *The Rise of Endymion* received the Locus Award in 1998.[8] All four novels were also nominated for various science fiction awards.

An event series is being developed by Bradley Cooper, Graham King, and Todd Phillips for Syfy based on the first novel *Hyperion*.[9]

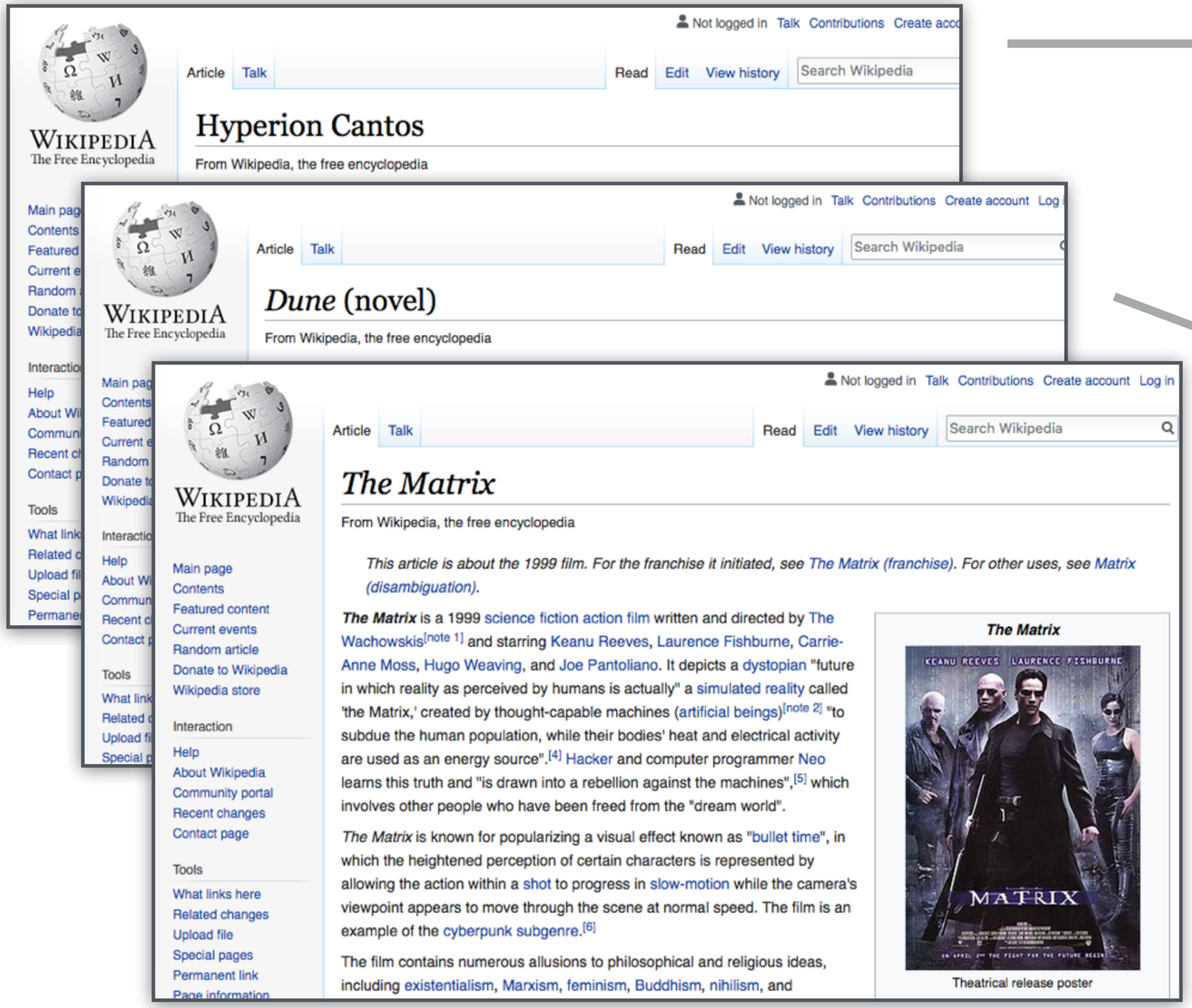
Dune is a 1965 science fiction novel by American author Frank Herbert, originally published as two separate serials in *Analog* magazine. It tied with Roger Zelazny's *This Immortal* for the Hugo Award in 1966,[3] and it won the inaugural Nebula Award for Best Novel.[4] It is the first installment of the *Dune* saga, and in 2003 was cited as the world's best-selling science fiction novel.[5][6]

Set in the distant future amidst a feudal interstellar society in which noble houses, in control of individual planets, owe allegiance to the Padishah Emperor, *Dune* tells the story of young Paul Atreides, whose noble family accepts the stewardship of the planet Arrakis. It is an inhospitable and sparsely populated desert wasteland, but is also the only source of melange, also known as "spice", a drug that enhances mental abilities. As melange is the most important and valuable substance in the universe, control of Arrakis is a coveted—and dangerous—undertaking. The story explores the multi-layered interactions of politics, religion, ecology, technology, and human emotion, as the factions of the empire confront each other in a struggle for the control of Arrakis

---

The Matrix is a 1999 science fiction action film written and directed by The Wachowskis[note 1] and starring Keanu Reeves, Laurence Fishburne, Carrie-Anne Moss, Hugo Weaving, and Joe Pantoliano. It depicts a dystopian "future in which reality as perceived by humans is actually" a simulated reality called 'the Matrix,' created by thought-capable machines (artificial beings)[note 2] "to subdue the human population, while their bodies' heat and electrical activity are used as an energy source".[4] Hacker and computer programmer Neo learns this truth and "is drawn into a rebellion against the machines",[5] which involves other people who have been freed from the "dream world".

The Matrix is known for popularizing a visual effect known as "bullet time", in which the heightened perception of certain characters is represented by allowing the action within a shot to progress in slow-motion while the camera's viewpoint appears to move through the scene at normal speed. The film is an example of the cyberpunk subgenre.[6]



The **Hyperion Cantos** is a series of science fiction novels by Dan Simmons. The title [illegible] in the series, *Hyperion* and *The Fall of Hyperion*,[1][2] and later came to refer to the overall storyline, including *Endymion*, *The Rise of Endymion*, and a number of short stories.[3][4] More narrowly, inside the fictional storyline, after the first volume, the Hyperion Cantos is an epic poem written by the character Martin Silenus covering in verse form the events of the first book.[5]

Of the four novels, *Hyperion* received the Hugo and Locus Awards in 1990;[6] *The Fall of Hyperion* won the Locus and British Science Fiction Association Awards in 1991;[7] and *The Rise of Endymion* received the Locus Award in 1998.[8] All four novels were also nominated for various science fiction awards.

An event series is being developed by Bradley Cooper, Graham King, and Todd Phillips for Syfy based on the first novel *Hyperion*.[9]

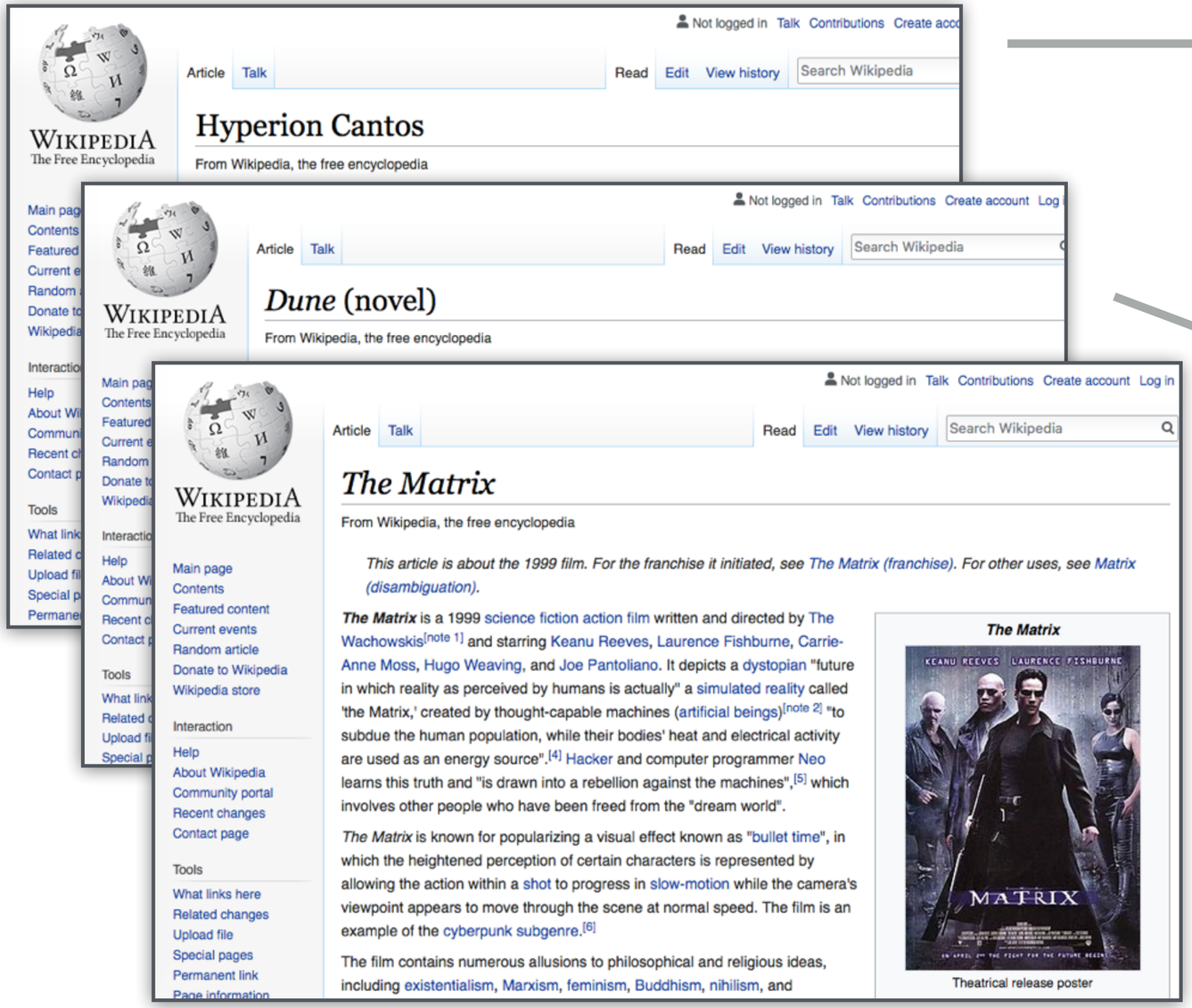
Dune is a 1965 science fiction novel by American author Frank Herbert, originally published as two separate serials in *Analog* magazine. It tied with Roger Zelazny's *This Immortal* for the Hugo Award in 1966,[3] and it won the inaugural Nebula Award for Best Novel.[4] It is the first installment of the *Dune* saga, and in 2003 was cited as the world's best-selling science fiction novel.[5][6]

Set in the distant future amidst a feudal interstellar society in which noble houses, in control of individual planets, owe allegiance to the Padishah Emperor, *Dune* tells the story of young Paul Atreides, whose noble family accepts the stewardship of the planet Arrakis. It is an inhospitable and sparsely populated desert wasteland, but is also the only source of melange, also known as "spice", a drug that enhances mental abilities. As melange is the most important and valuable substance in the universe, control of Arrakis is a coveted—and dangerous—undertaking. The story explores the multi-layered interactions of politics, religion, ecology, technology, and human emotion, as the factions of the empire confront each other in a struggle for the control of Arrakis

---

The Matrix is a 1999 science fiction action film written and directed by The Wachowskis[note 1] and starring Keanu Reeves, Laurence Fishburne, Carrie-Anne Moss, Hugo Weaving, and Joe Pantoliano. It depicts a dystopian "future in which reality as perceived by humans is actually" a simulated reality called 'the Matrix,' created by thought-capable machines (artificial beings)[note 2] "to subdue the human population, while their bodies' heat and electrical activity are used as an energy source".[4] Hacker and computer programmer Neo learns this truth and "is drawn into a rebellion against the machines",[5] which involves other people who have been freed from the "dream world".

The Matrix is known for popularizing a visual effect known as "bullet time", in which the heightened perception of certain characters is represented by allowing the action within a shot to progress in slow-motion while the camera's viewpoint appears to move through the scene at normal speed. The film is an example of the cyberpunk subgenre.[6]



The **Hyperion Cantos** is a series of science fiction novels by Dan Simmons. The title was originally used for the collection of the first pair of books in the series, *Hyperion* and *The Fall of Hyperion*,[1][2] and later came to refer to the overall storyline, including *Endymion*, *The Rise of Endymion*, and a number of short stories.[3][4] More narrowly, inside the fictional storyline, after the first volume, the Hyperion Cantos is an epic poem written by the character Martin Silenus covering in verse form the events of the first book.[5]

Of the four novels, *Hyperion* received the Hugo and Locus Awards in 1990;[6] *The Fall of Hyperion* won the Locus and British Science Fiction Association Awards in 1991;[7] and *The Rise of Endymion* received the Locus Award in 1998.[8] All four novels were also nominated for various science fiction awards.

An event series is being developed by Bradley Cooper, Graham King, and Todd Phillips for Syfy based on the first novel *Hyperion*.[9]

Dune is a 1965 science fiction novel by American author Frank Herbert, [illegible]ine. It tied with Roger Zelazny's *This Immortal* for the Hugo Award in 1966,[3] and it won the inaugural Nebula Award for Best Novel.[4] It is the first installment of the *Dune* saga, and in 2003 was cited as the world's best-selling science fiction novel.[5][6]

Set in the distant future amidst a feudal interstellar society in which noble houses, in control of individual planets, owe allegiance to the Padishah Emperor, *Dune* tells the story of young Paul Atreides, whose noble family accepts the stewardship of the planet Arrakis. It is an inhospitable and sparsely populated desert wasteland, but is also the only source of melange, also known as "spice", a drug that enhances mental abilities. As melange is the most important and valuable substance in the universe, control of Arrakis is a coveted—and dangerous—undertaking. The story explores the multi-layered interactions of politics, religion, ecology, technology, and human emotion, as the factions of the empire confront each other in a struggle for the control of Arrakis

The Matrix is a 1999 science fiction action film written and directed by The Wachowskis[note 1] and starring Keanu Reeves, Laurence Fishburne, Carrie-Anne Moss, Hugo Weaving, and Joe Pantoliano. It depicts a dystopian "future in which reality as perceived by humans is actually" a simulated reality called 'the Matrix,' created by thought-capable machines (artificial beings)[note 2] "to subdue the human population, while their bodies' heat and electrical activity are used as an energy source".[4] Hacker and computer programmer Neo learns this truth and "is drawn into a rebellion against the machines",[5] which involves other people who have been freed from the "dream world".

The Matrix is known for popularizing a visual effect known as "bullet time", in which the heightened perception of certain characters is represented by allowing the action within a shot to progress in slow-motion while the camera's viewpoint appears to move through the scene at normal speed. The film is an example of the cyberpunk subgenre.[6]

“Thou shalt not make a machine in the likeness of a human mind”
~Dune

Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

thou	shalt	not	make	a	machine	in	the	…

Dataset

input 1	input 2	output

Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

thou	shalt	not	make	a	machine	in	the	…

Dataset

input 1	input 2	output

Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

thou	shalt	not	make	a	machine	in	the	…

Dataset

input 1	input 2	output
thou	shalt	not

Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

thou	shalt	not	make	a	machine	in	the	…
thou	shalt	not	make	a	machine	in	the	

Dataset

input 1	input 2	output
thou	shalt	not

Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

thou	shalt	not	make	a	machine	in	the	…
thou	shalt	not	make	a	machine	in	the	

Dataset

input 1	input 2	output
thou	shalt	not
shalt	not	make

Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

thou	shalt	not	make	a	machine	in	the	…
thou	shalt	not	make	a	machine	in	the	
thou	shalt	not	make	a	machine	in	the	

Dataset

input 1	input 2	output
thou	shalt	not
shalt	not	make
not	make	a

Thou shalt not make a machine in the likeness of a human mind

Sliding window across running text

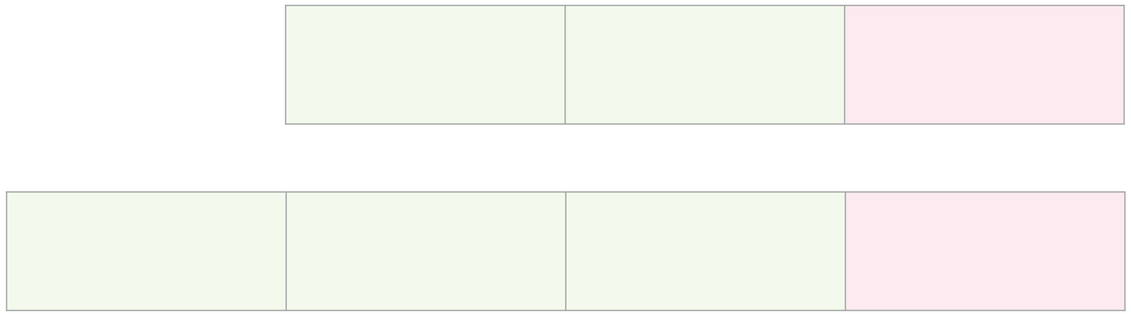
thou	shalt	not	make	a	machine	in	the	…
thou	shalt	not	make	a	machine	in	the	
thou	shalt	not	make	a	machine	in	the	
thou	shalt	not	make	a	machine	in	the	
thou	shalt	not	make	a	machine	in	the	

Dataset

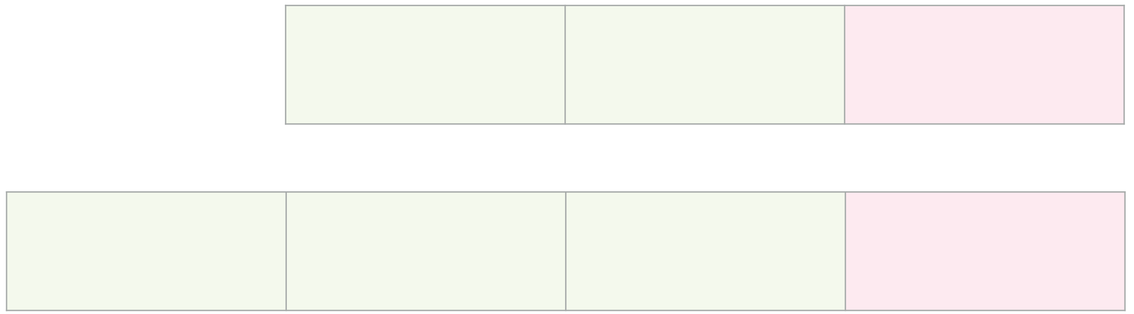
input 1	input 2	output
thou	shalt	not
shalt	not	make
not	make	a
make	a	machine
a	machine	in

Jay was hit by a _____

Jay was hit by a _____



Jay was hit by a bus?



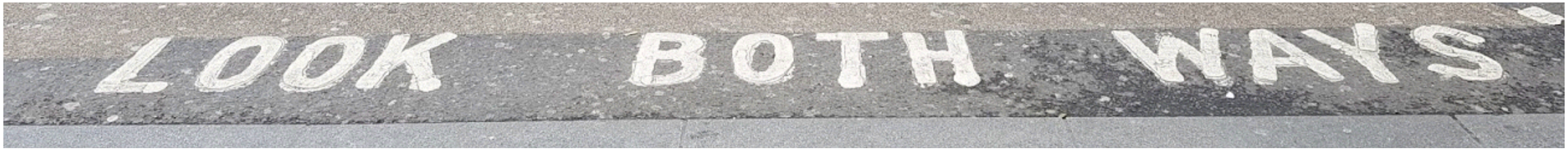
Jay was hit by a _____ bus in...

~~bus?~~

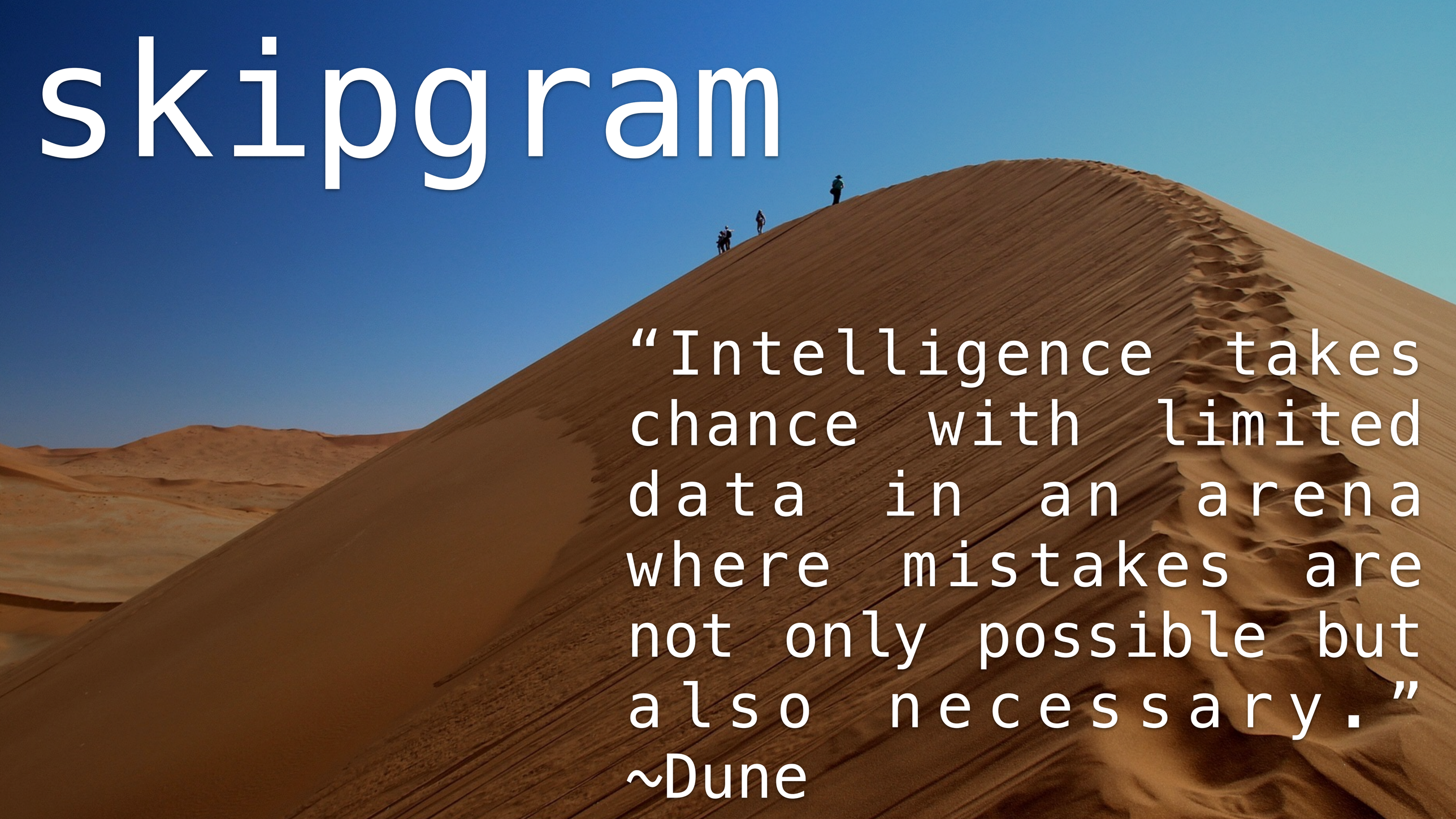
Jay was hit by a red bus in...

Jay was hit by a _____

Jay was hit by a _____ bus in…



skipgram



“Intelligence takes chance with limited data in an arena where mistakes are not only possible but also necessary.”
~Dune

Jay was hit by a _____ bus in...

Jay was hit by a _____ bus in...

by	a	red	bus	in

input 1	input 2	input 3	input 4	output
by	a	bus	in	red

Continuous Bag of Words (CBoW)

Jay was hit by a red bus in...

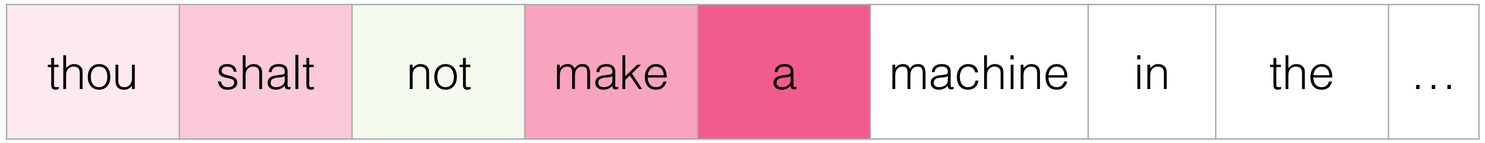
Jay was hit by a red bus in...

Jay was hit by a red bus in...

by	a	red	bus	in

input	output
red	by
red	a
red	bus
red	in

Thou shalt not make a machine in the likeness of a human mind



input word	target word

Thou shalt not make a machine in the likeness of a human mind

thou	shalt	not	make	a	machine	in	the	…

input word	target word
not	thou
not	shalt
not	make
not	a

Thou shalt not make a machine in the likeness of a human mind

thou	shalt	not	make	a	machine	in	the	…

thou	shalt	not	make	a	machine	in	the	…

input word	target word
not	thou
not	shalt
not	make
not	a

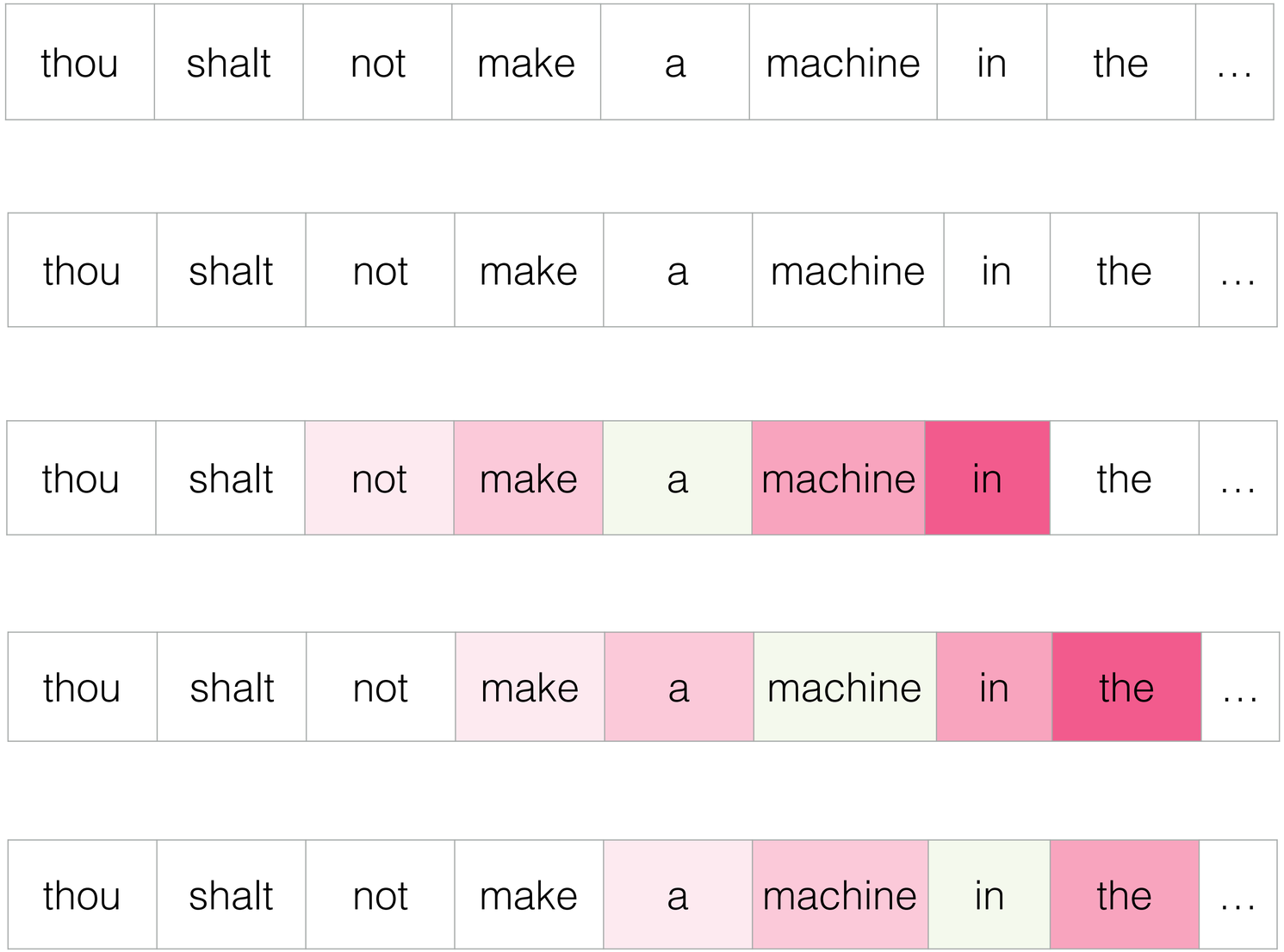
Thou shalt not make a machine in the likeness of a human mind

thou	shalt	not	make	a	machine	in	the	…

thou	shalt	not	make	a	machine	in	the	…

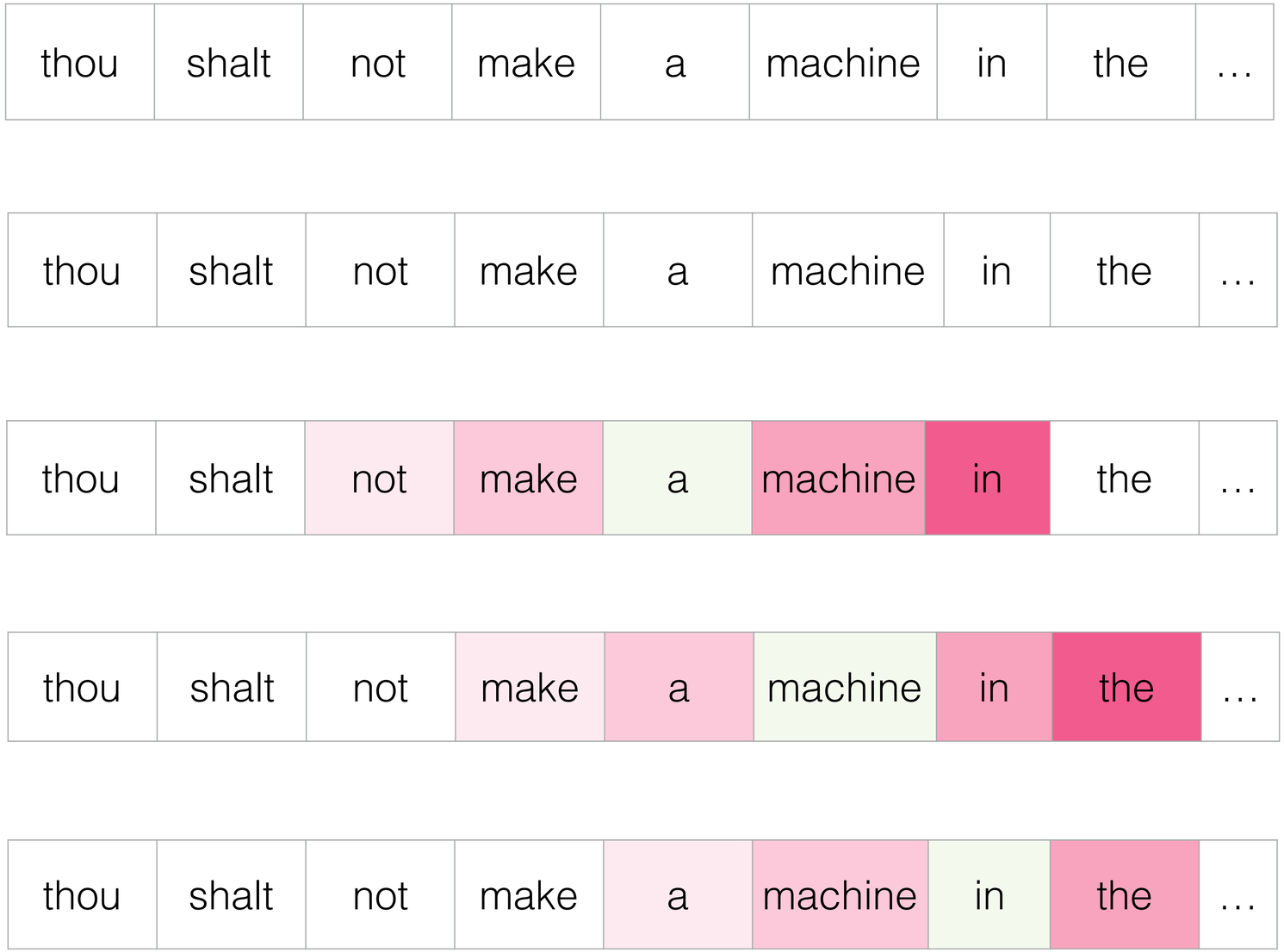
input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine

Thou shalt not make a machine in the likeness of a human mind



input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine

Thou shalt not make a machine in the likeness of a human mind

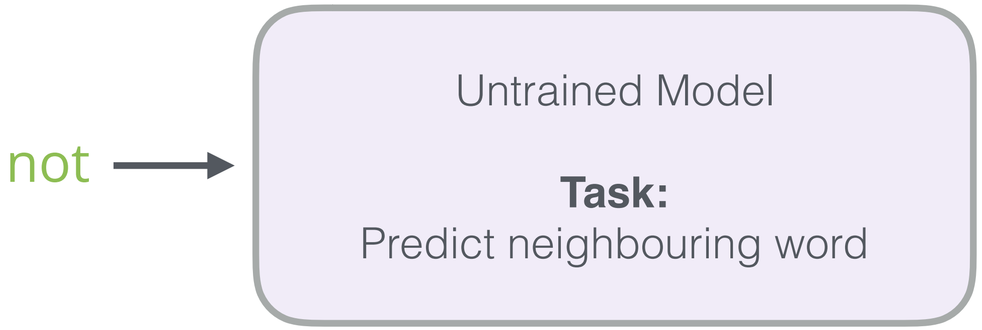


input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

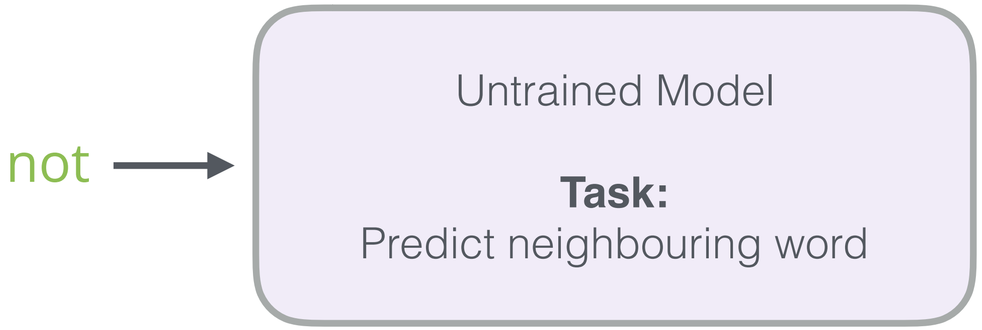
input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness



input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

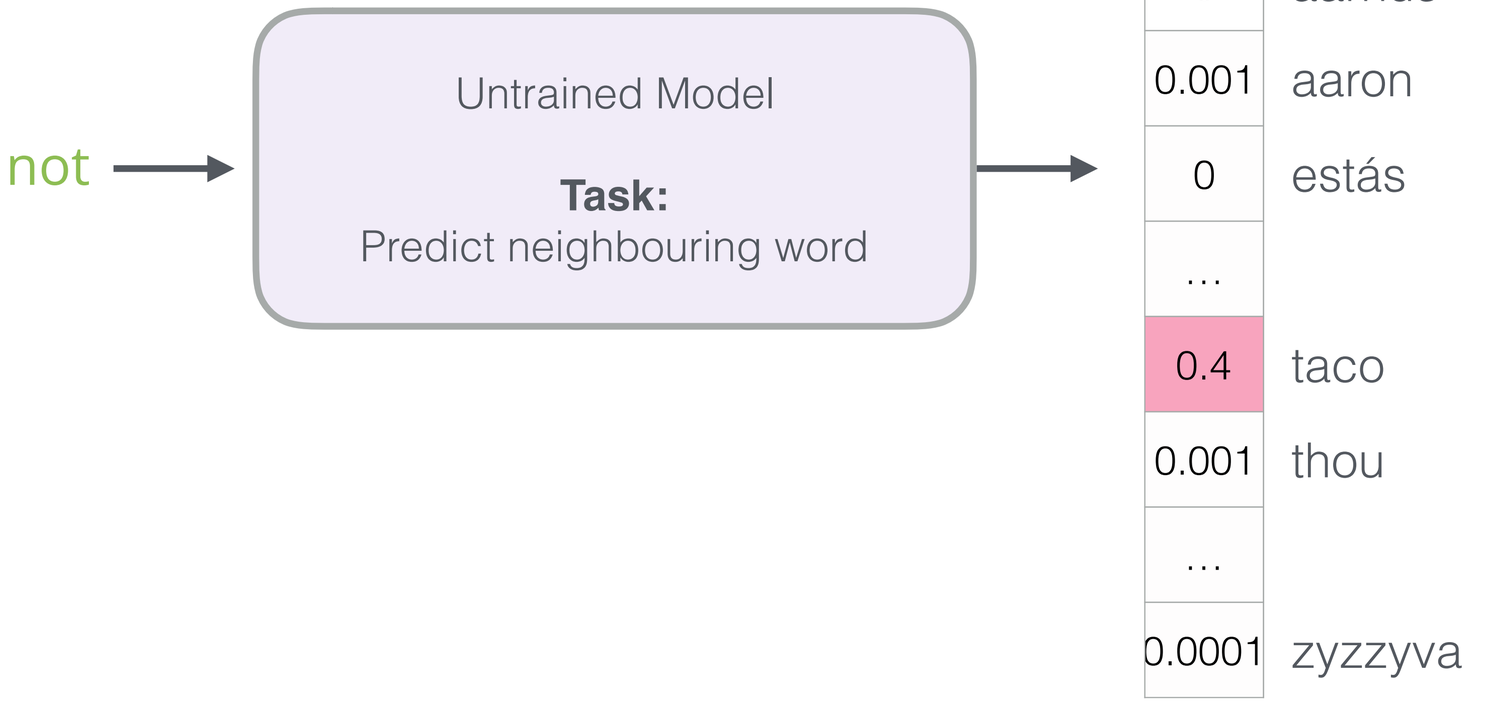


1) Look up embeddings

2) Calculate Prediction

3) Project to Output Vocabulary

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness



input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

Model Prediction

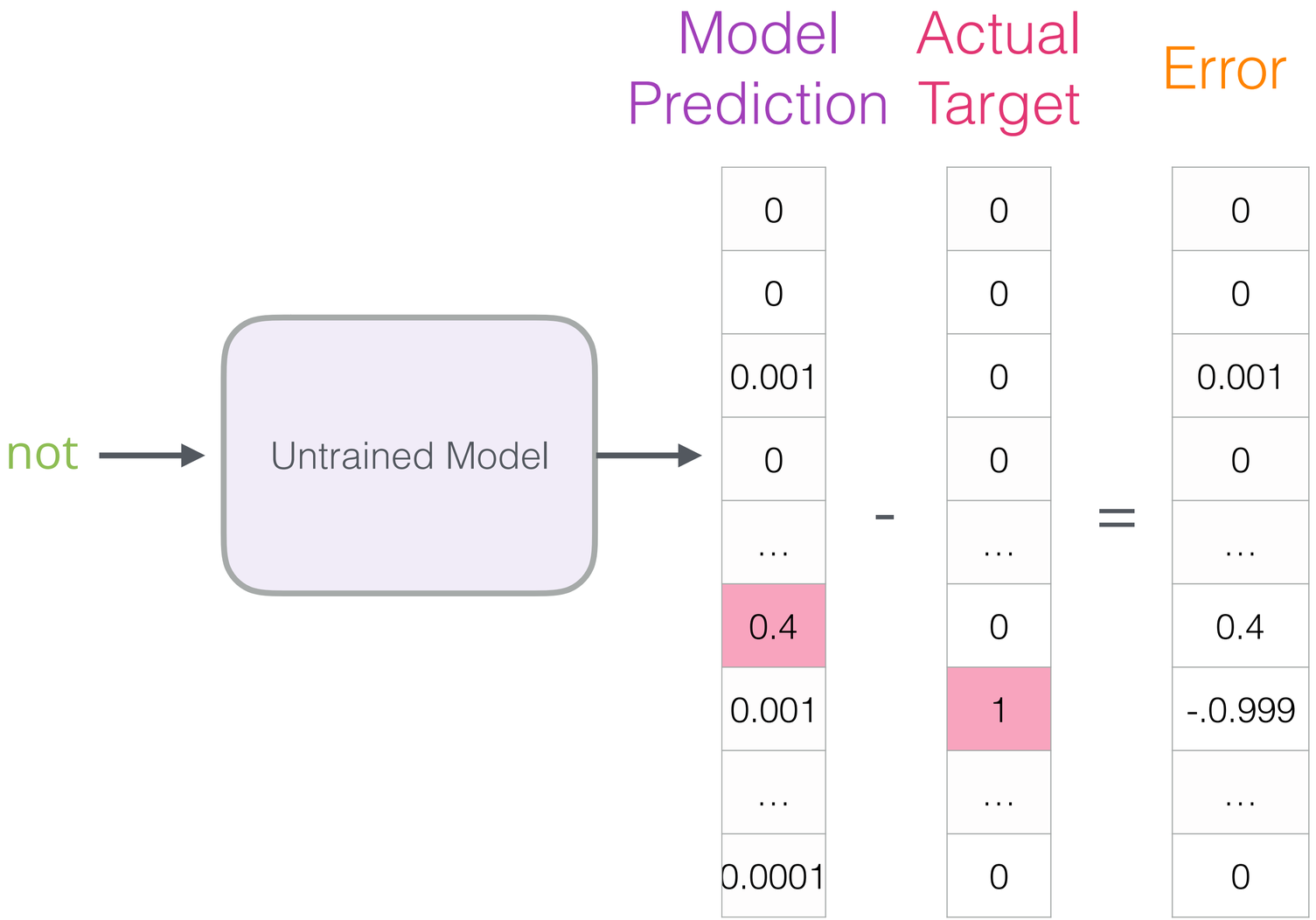
0
0
0.001
0
…
0.4
0.001
…
0.0001

-

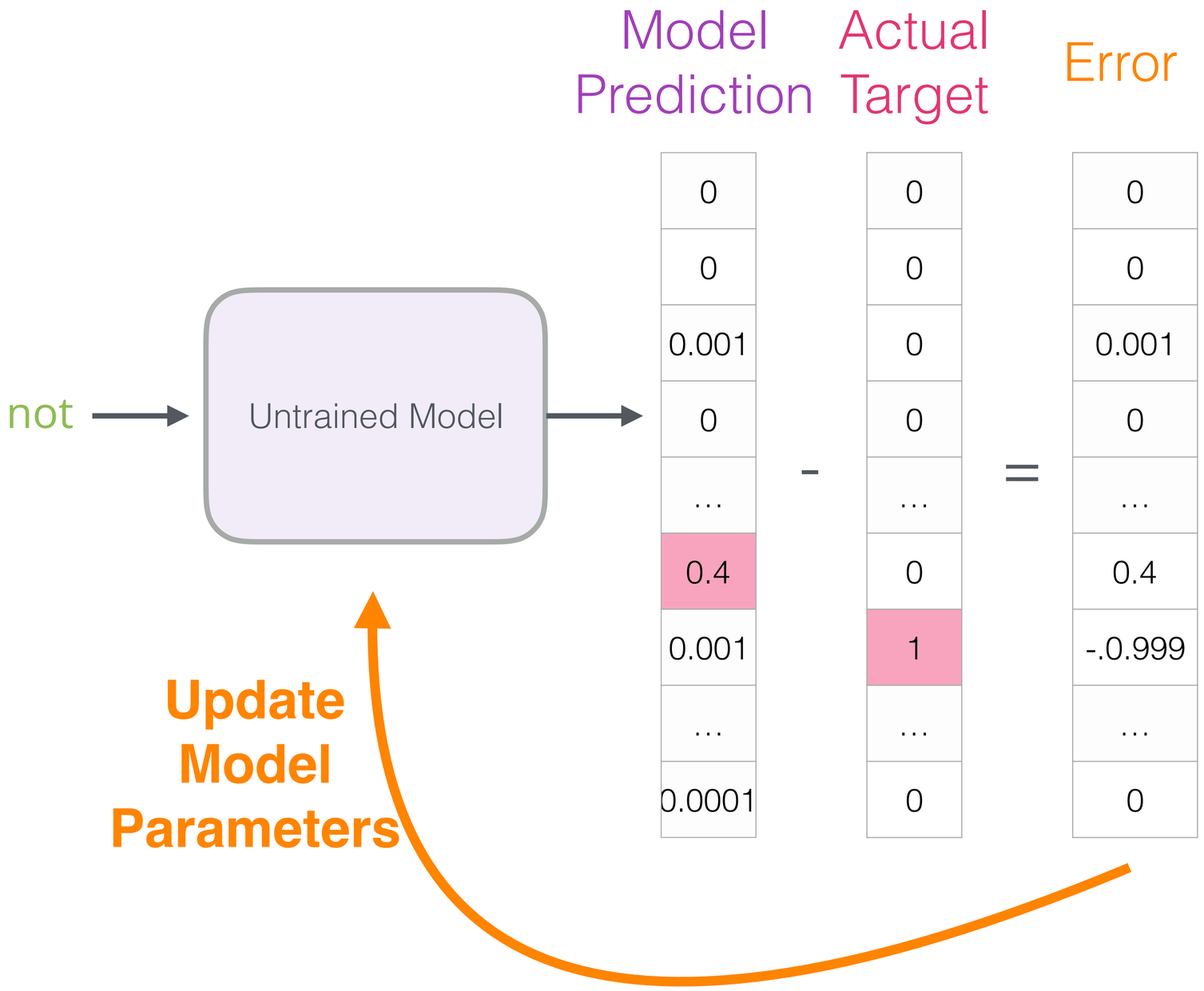
Actual Target

0	aardvark
0	aarhus
0	aaron
0	estás
…	
0	taco
1	thou
…	
0	zyzzyva

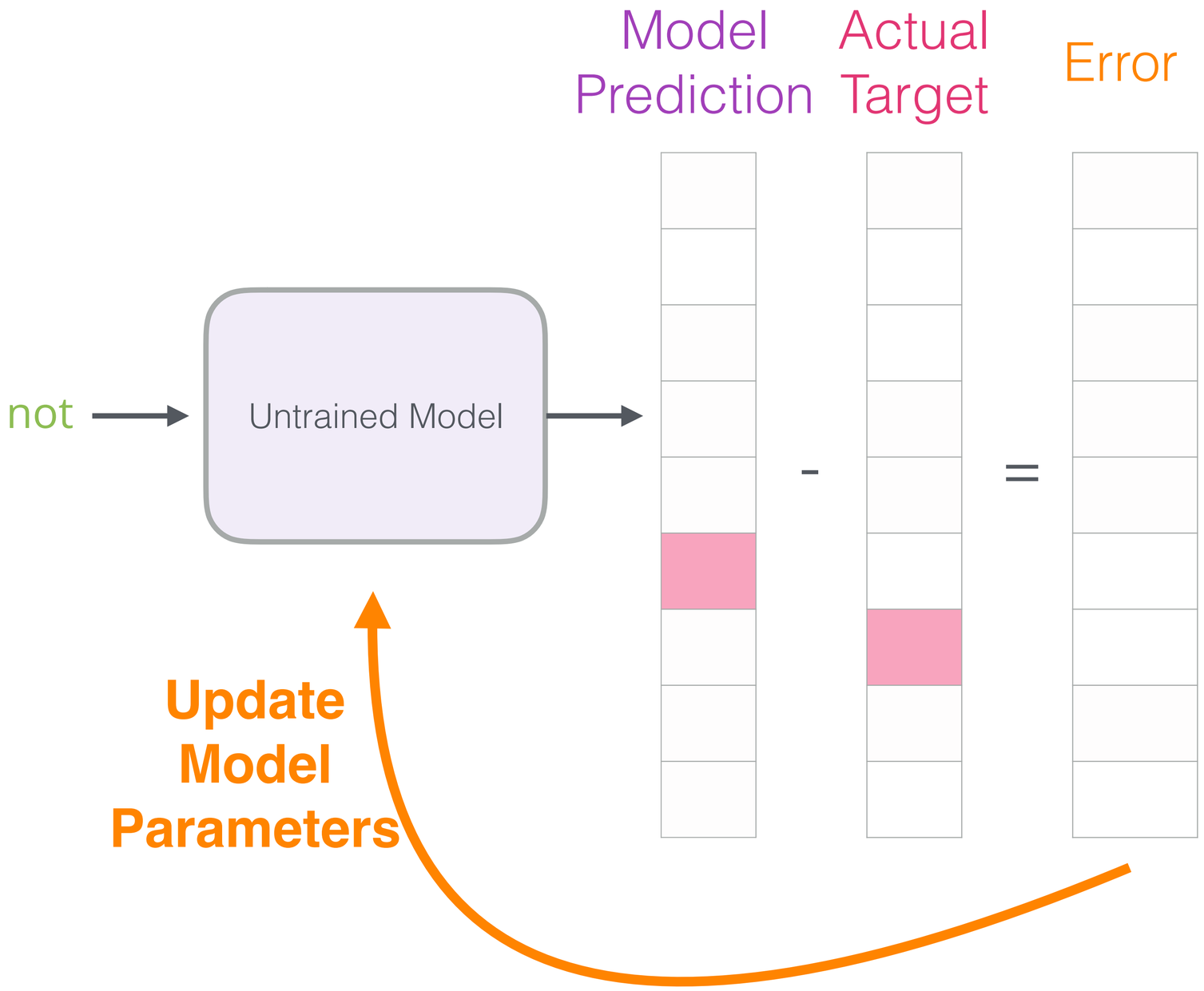
input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness



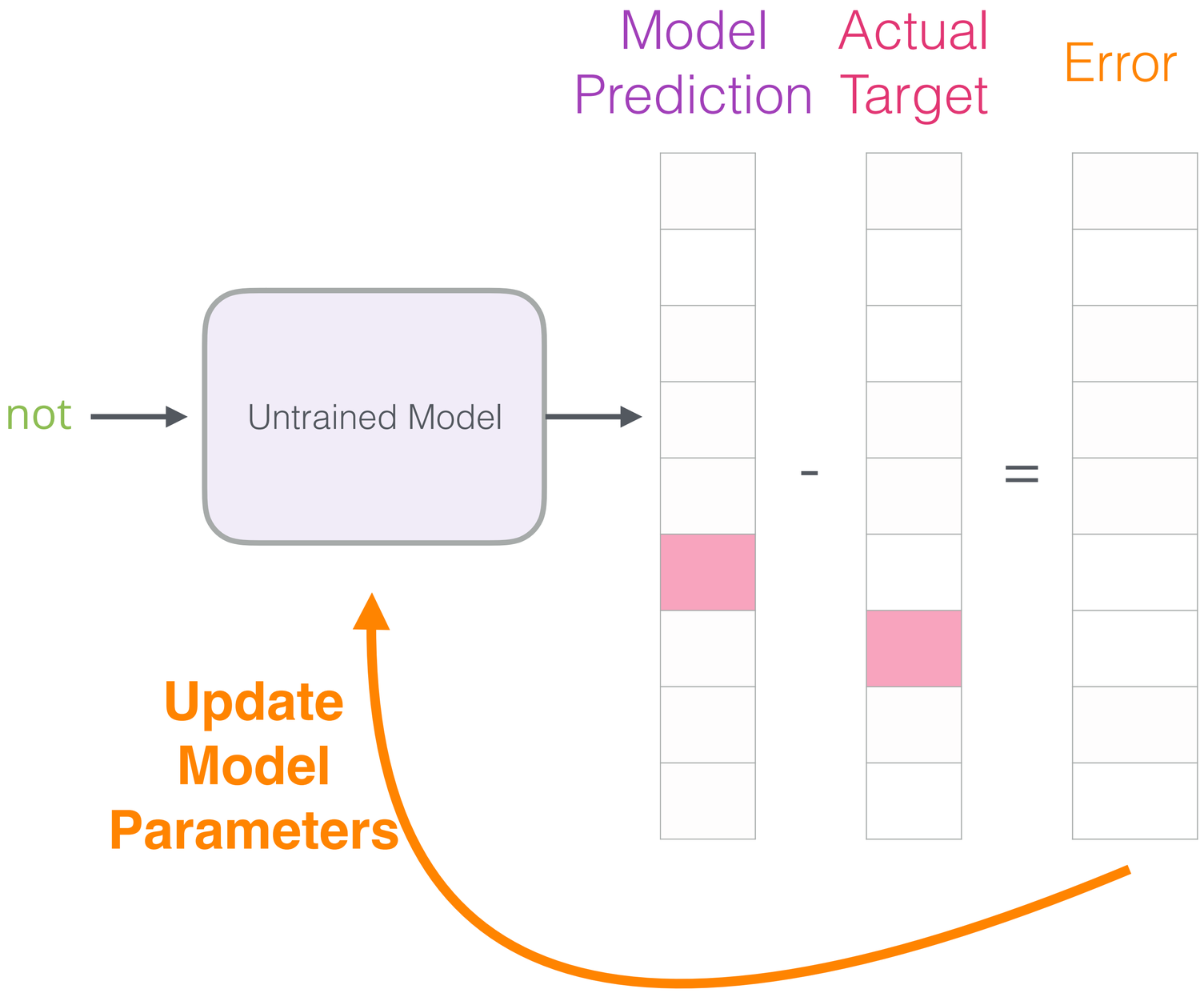
input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness



input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness



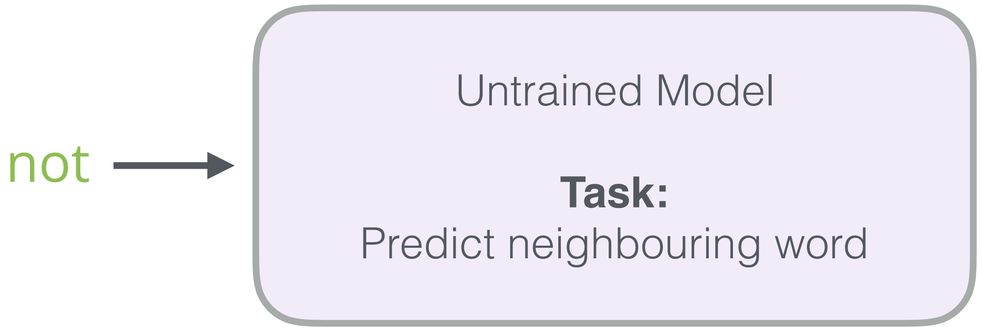
input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness



Negative Sampling

“To attempt an understanding of Muad'Dib without understanding his mortal enemies, the Harkonnens, is to attempt seeing Truth without knowing Falsehood. It is the attempt to see the Light without knowing Darkness. It cannot be.”
~Dune

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

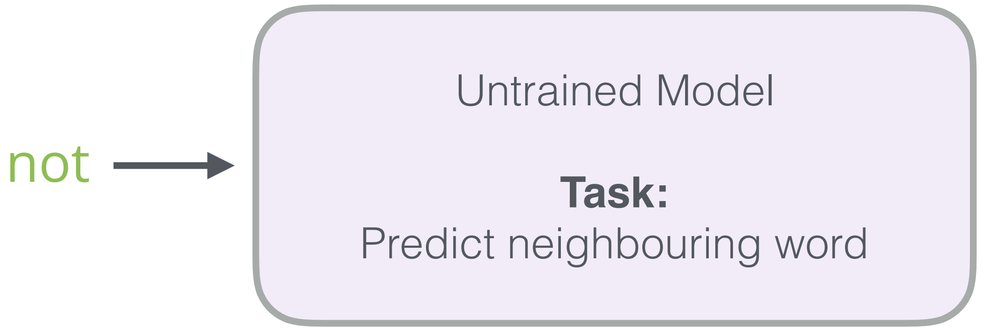


1) Look up embeddings

2) Calculate Prediction

3) Project to Output Vocabulary

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness



1) Look up embeddings

2) Calculate Prediction

3) Project to Output Vocabulary

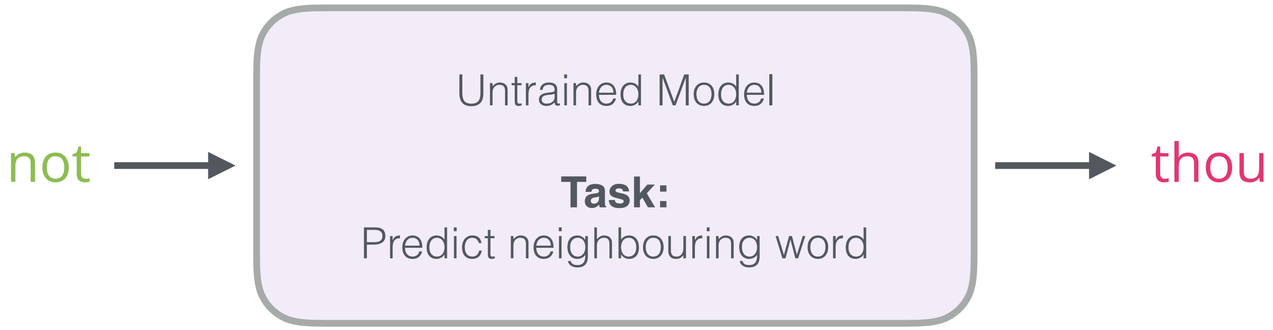
[Computationally Intensive]

1- Generate high quality word embeddings
(Don’t worry about next-word prediction)

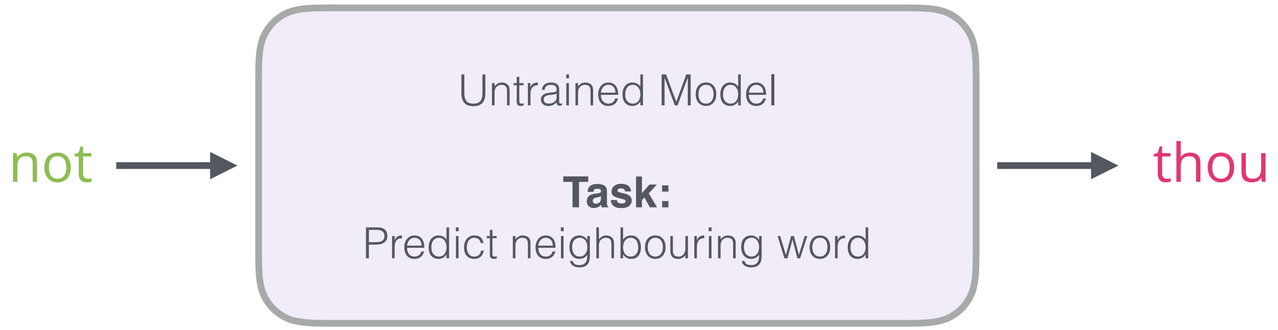
2- Use these high-quality embeddings to train a language model (to do next-word prediction)

We will not worry about this in this talk

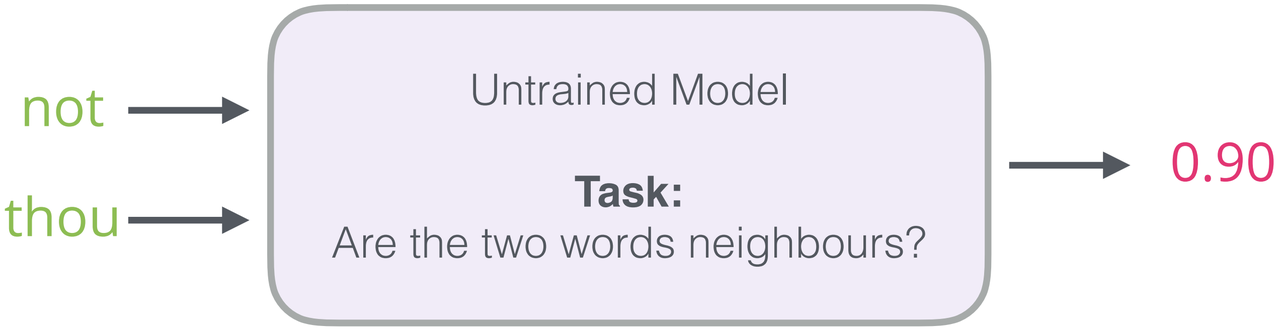
Change Task from



Change Task from



To:



Noise-contrastive estimation: A new estimation principle for unnormalized statistical models

Michael Gutmann
Dept of Computer Science
and HIIT, University of Helsinki
michael.gutmann@helsinki.fi

Aapo Hyvärinen
Dept of Mathematics & Statistics, Dept of Computer
Science and HIIT, University of Helsinki
aapo.hyvarinen@helsinki.fi

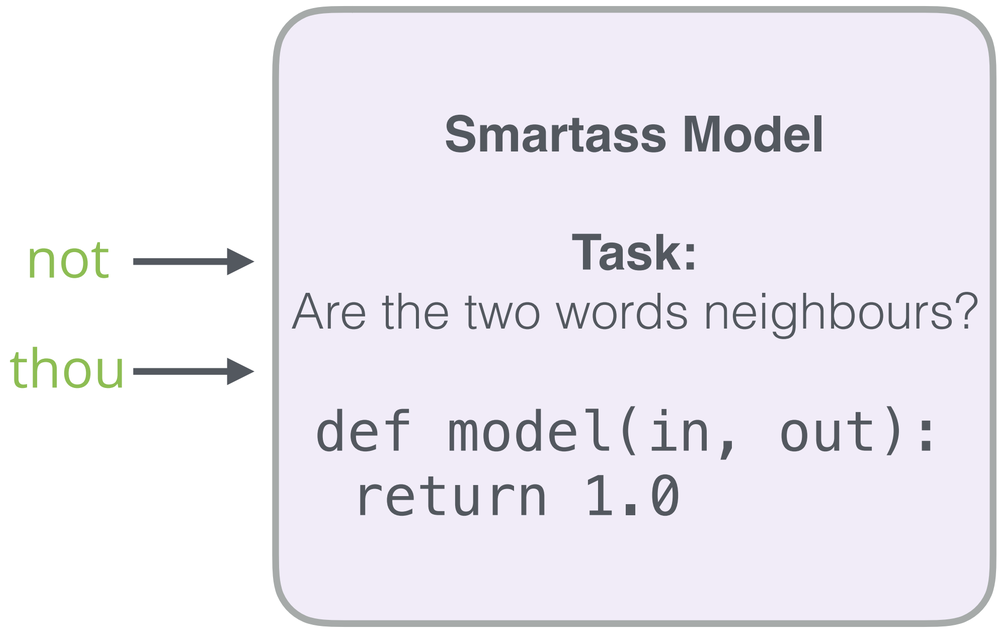
http://proceedings.mlr.press/v9/gutmann10a/gutmann10a.pdf

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine

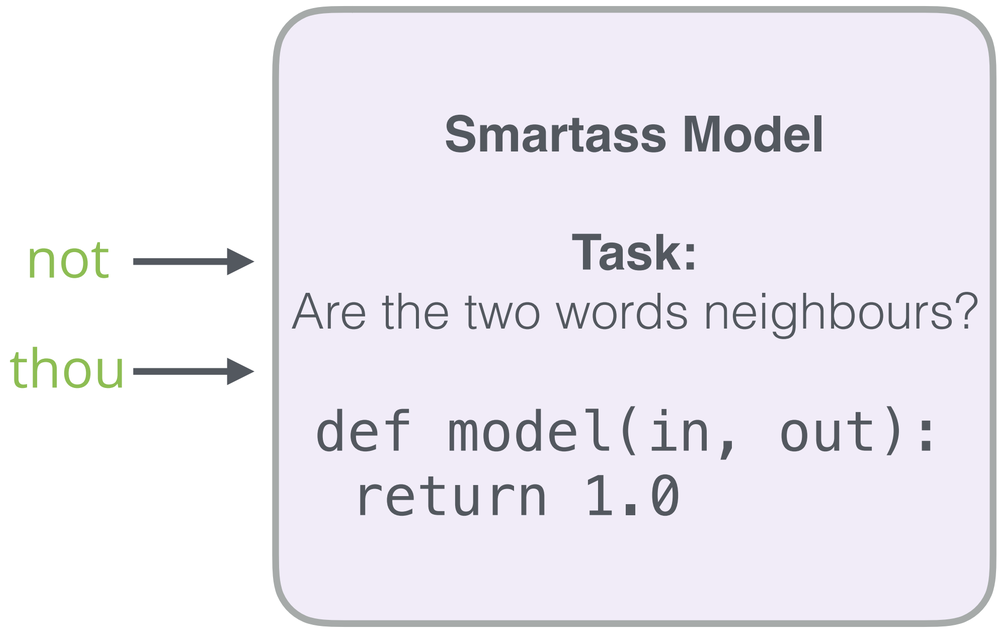
input word	output word	target
not	thou	**1**
not	shalt	**1**
not	make	**1**
not	a	**1**
make	shalt	**1**
make	not	**1**
make	a	**1**
make	machine	**1**

input word	output word	target
not	thou	**1**
not	shalt	**1**
not	make	**1**
not	a	**1**
make	shalt	**1**
make	not	**1**
make	a	**1**
make	machine	**1**

input word	output word	target
not	thou	1
not	shalt	1
not	make	1
not	a	1
make	shalt	1
make	not	1
make	a	1
make	machine	1



input word	output word	target
not	thou	1
not	shalt	1
not	make	1
not	a	1
make	shalt	1
make	not	1
make	a	1
make	machine	1



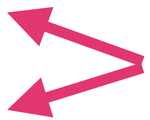
100% accuracy!!!

(but terrible embeddings)

input word	output word	target
not	thou	**1**
not	shalt	**1**
not	make	**1**
not	a	**1**
make	shalt	**1**
make	not	**1**
make	a	**1**
make	machine	**1**

input word	output word	target
not	thou	**1**
not	shalt	**1**
not	make	**1**

input word	output word	target
not	thou	**1**
not		**0**
not		**0**
not	shalt	**1**
not	make	**1**



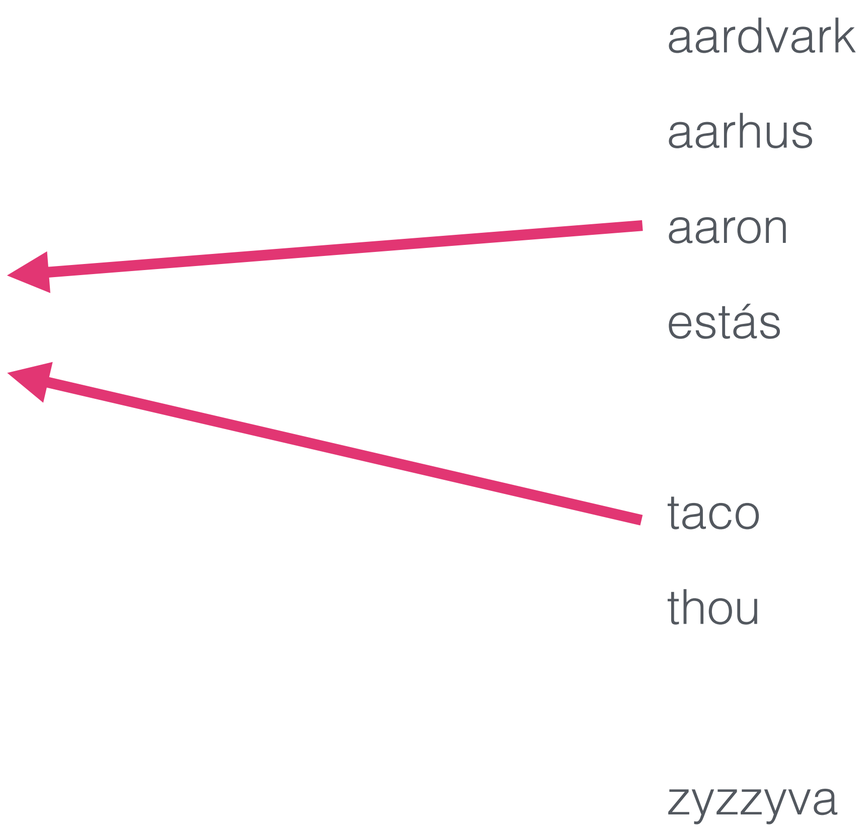
Negative examples

input word	output word	target
not	thou	**1**
not	?	**0**
not		**0**
not	shalt	**1**
not	make	**1**

Negative examples

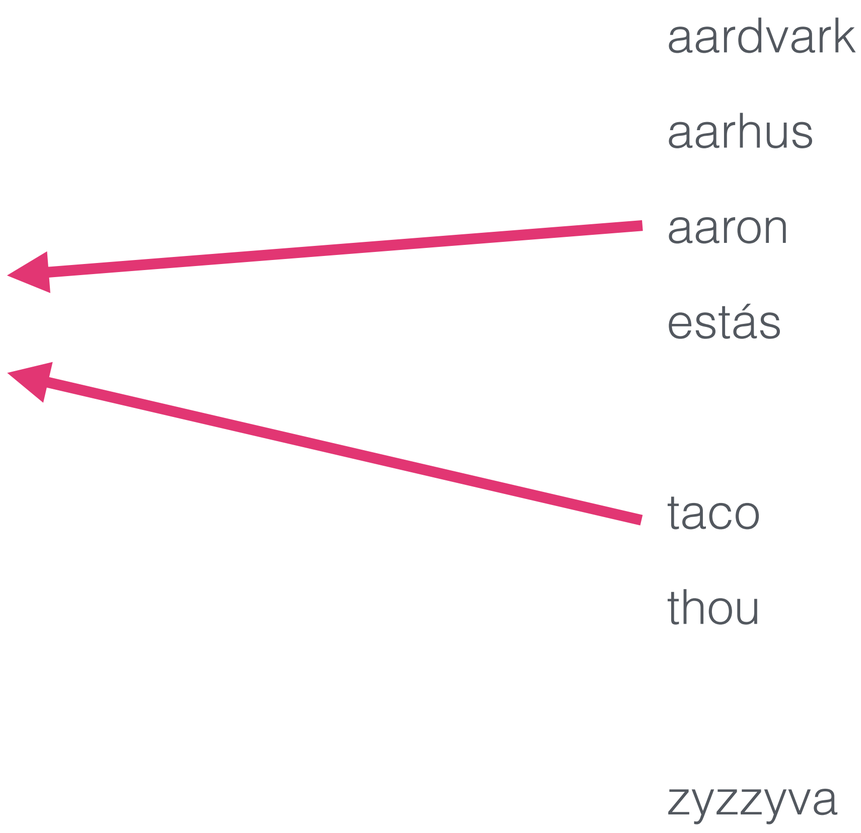
Pick randomly from vocabulary
(random sampling)

input word	output word	target
not	thou	**1**
not	?	**0**
not		**0**
not	shalt	**1**
not	make	**1**



Pick randomly from vocabulary
(random sampling)

input word	output word	target
not	thou	**1**
not	aaron	**0**
not	taco	**0**
not	shalt	**1**
not	make	**1**



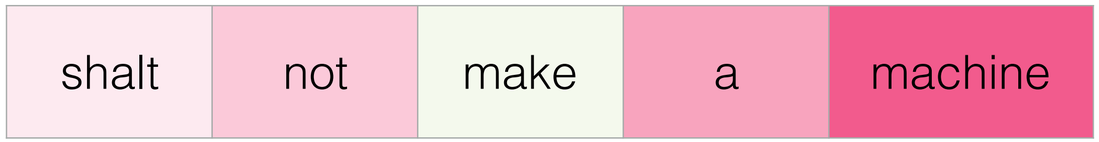
Pick randomly from vocabulary
(random sampling)

input word	output word	target
not	thou	**1**
not	aaron	**0**
not	taco	**0**
not	shalt	**1**
not	make	**1**

Word	Count	Probability
aardvark		
aarhus		
aaron		
estás		
taco		
thou		
zyzzyva		

Welcome to Word2vec

skipgram



Negative Sampling

input word	output word	target
not	thou	**1**
not	aaron	**0**
not	taco	**0**

thou	shalt	not	make	a	machine	in	the	…

homepage

`listing #104`

Site Search

`listing #73`

`listing #3065`

`listing #104` | `listing #73` | `listing #3065`

homepage

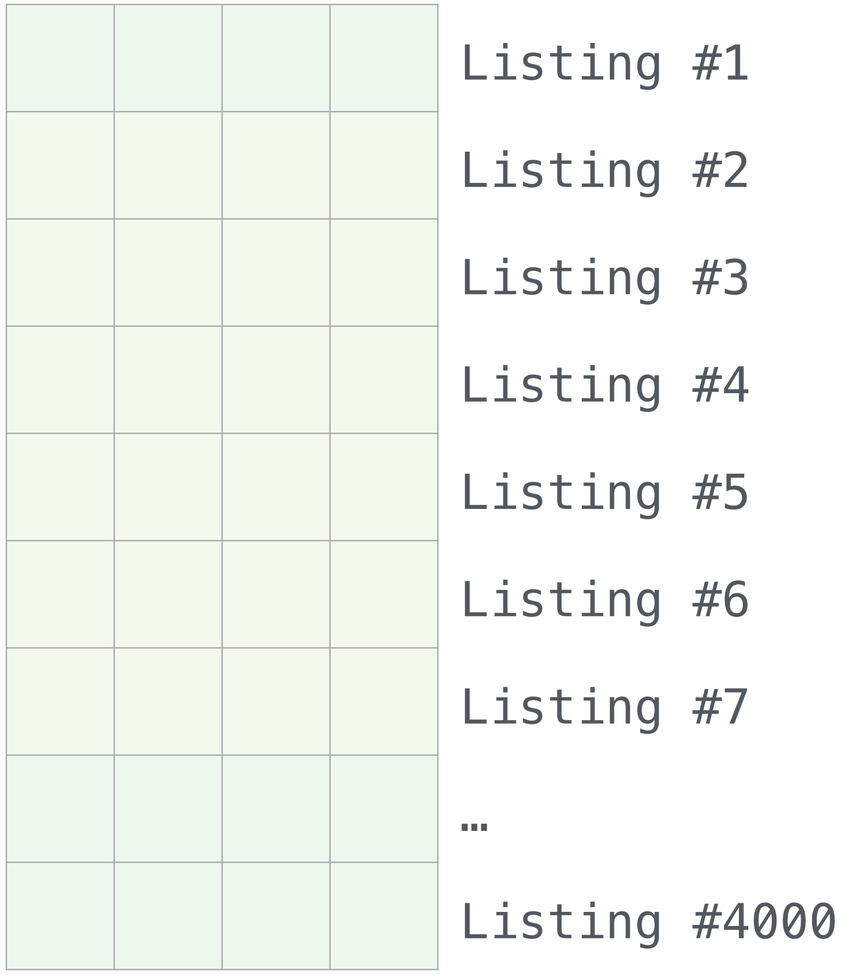
listing #104

Site Search

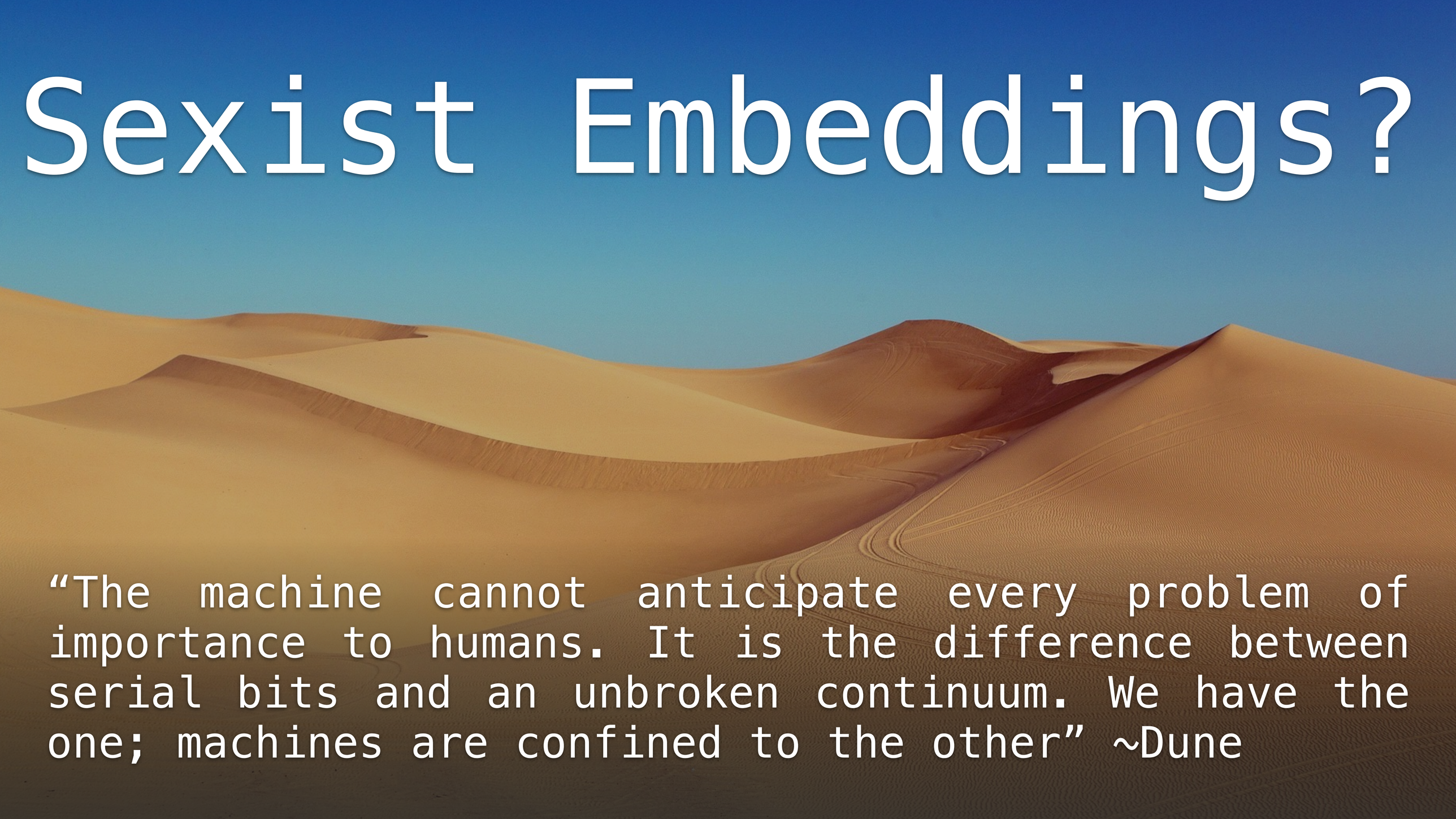
listing #73

listing #3065

listing #104	**listing #73**	**listing #3065**



Sexist Embeddings?



“The machine cannot anticipate every problem of importance to humans. It is the difference between serial bits and an unbroken continuum. We have the one; machines are confined to the other” ~Dune

france is to paris as italy is to rome

france is to paris as italy is to rome

man is to doctor as woman is to _____

man is to doctor as woman is to Nurse

```
model.most_similar(positive=["doctor", "woman"], negative=["man"])
```

```
[('nurse', 0.8404642343521118),
 ('child', 0.7663260102272034),
 ('pregnant', 0.7570130228996277),
 ('mother', 0.7517458200454712),
 ('patient', 0.7516663074493408),
 ('physician', 0.7507281303405762),
 ('dentist', 0.7360343933105469),
 ('therapist', 0.7342536449432373),
 ('parents', 0.7286345958709717),
 ('surgeon', 0.7165213227272034)]
```

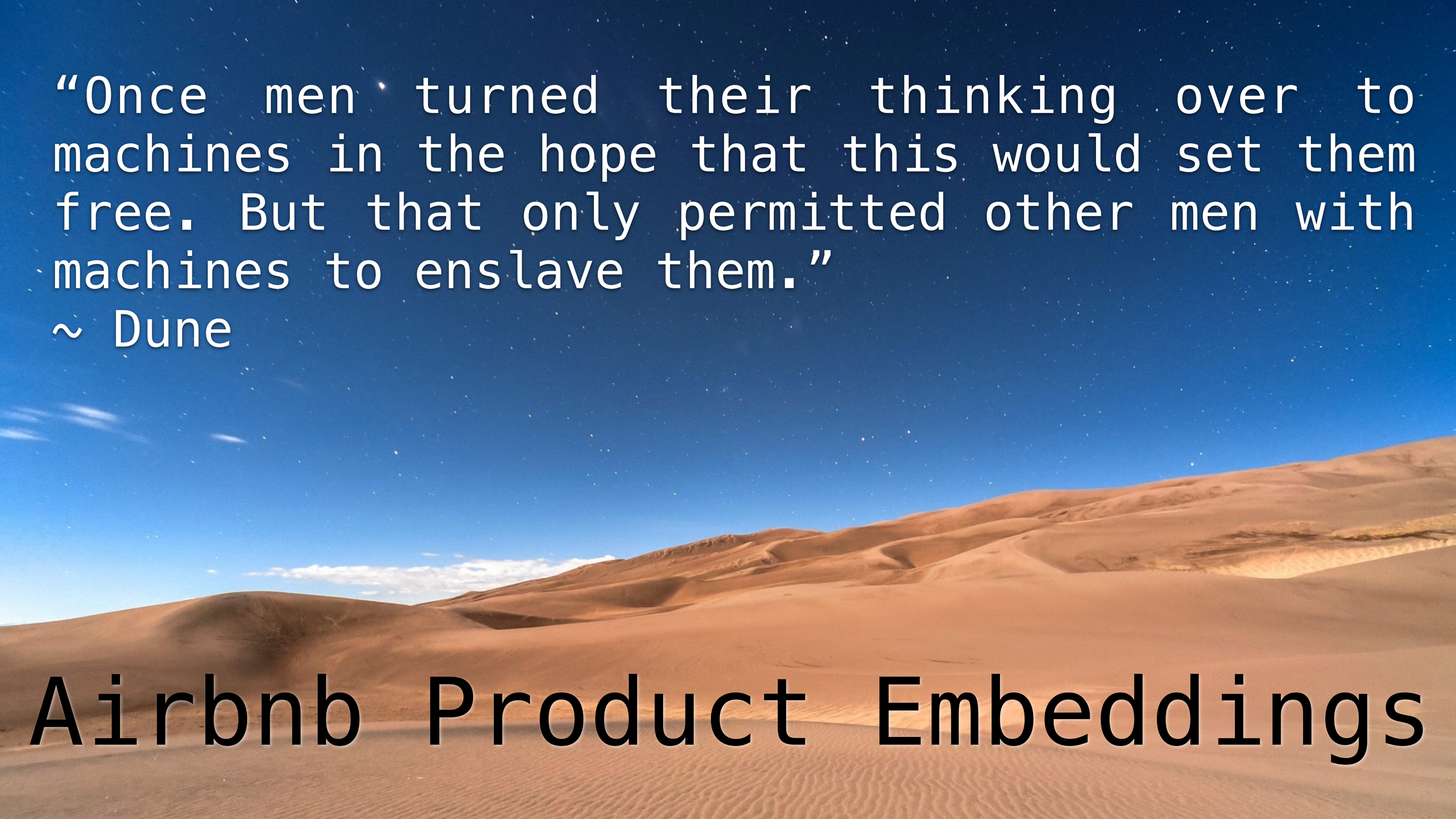
Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings

Tolga Bolukbasi[1], Kai-Wei Chang[2], James Zou[2], Venkatesh Saligrama[1,2], Adam Kalai[2]

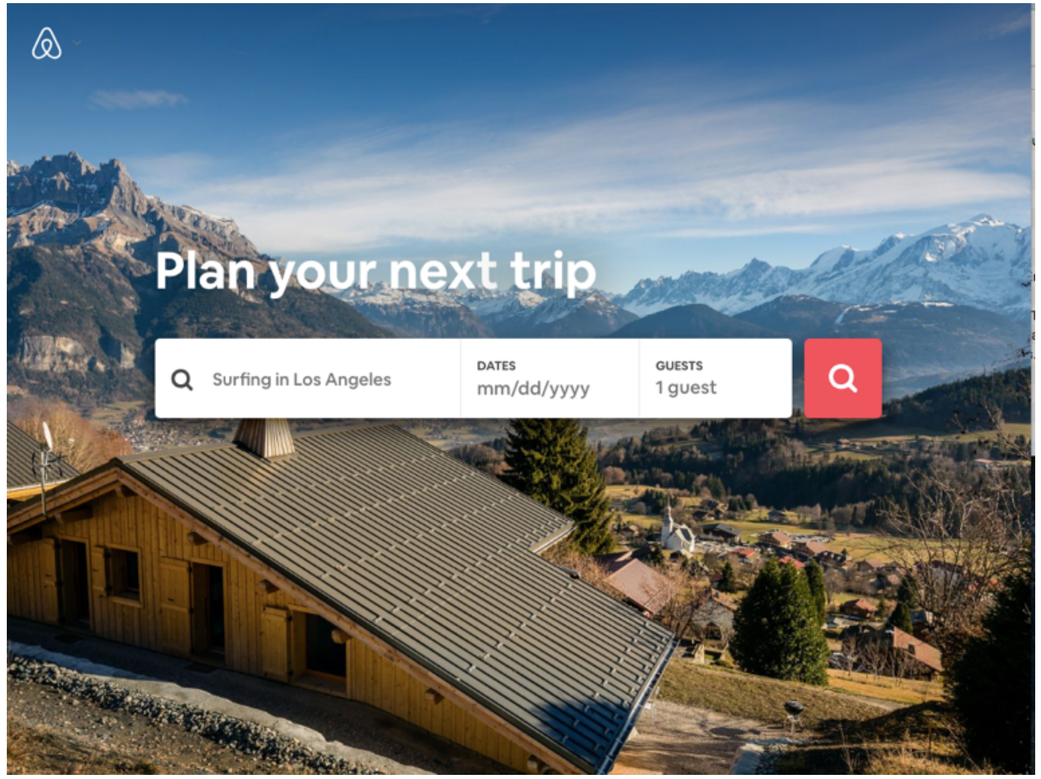
[1]Boston University, 8 Saint Mary's Street, Boston, MA

[2]Microsoft Research New England, 1 Memorial Drive, Cambridge, MA

tolgab@bu.edu, kw@kwchang.net, jamesyzou@gmail.com, srv@bu.edu, adam.kalai@microsoft.com

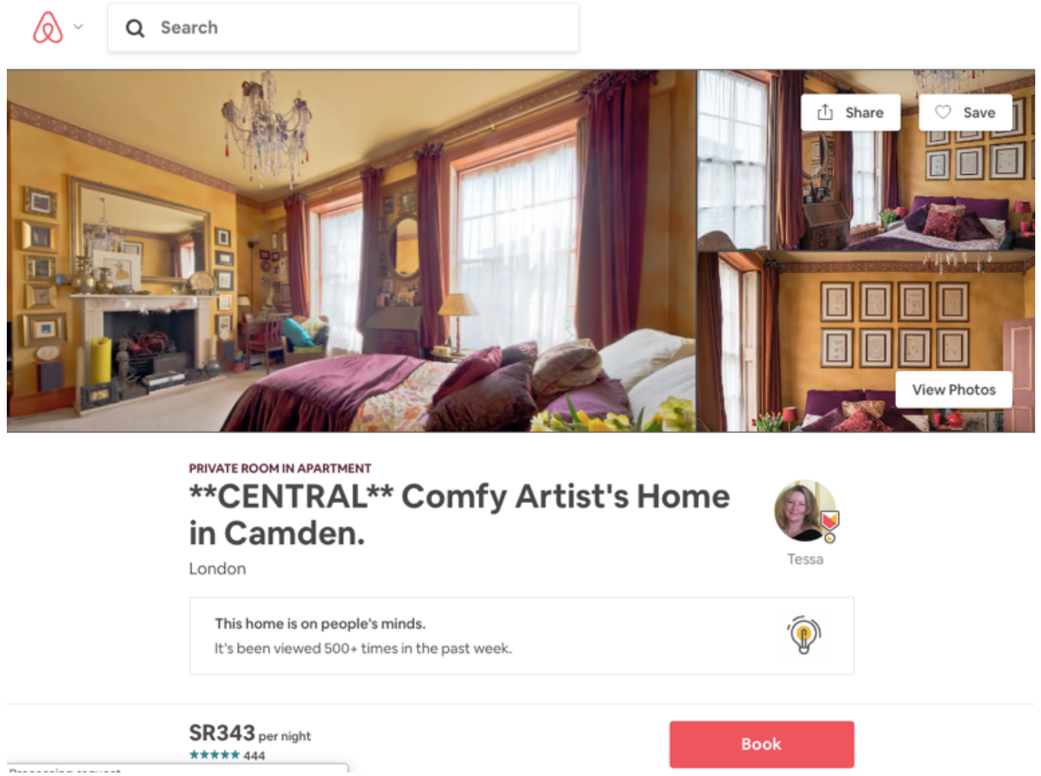

“Once men turned their thinking over to machines in the hope that this would set them free. But that only permitted other men with machines to enslave them.”
~ Dune
Airbnb Product Embeddings

homepage


Plan your next trip
Surfing in Los Angeles
DATES
mm/dd/yyyy
GUESTS
1 guest

homepage

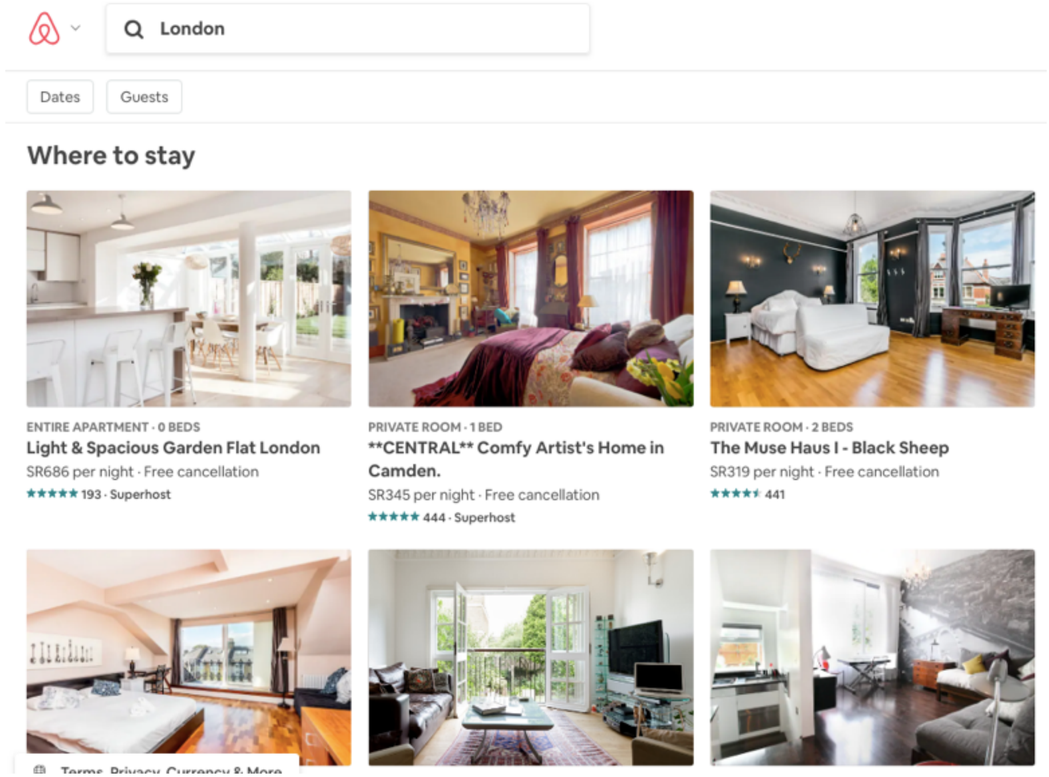
`listing #104`



homepage

listing #104

Site Search

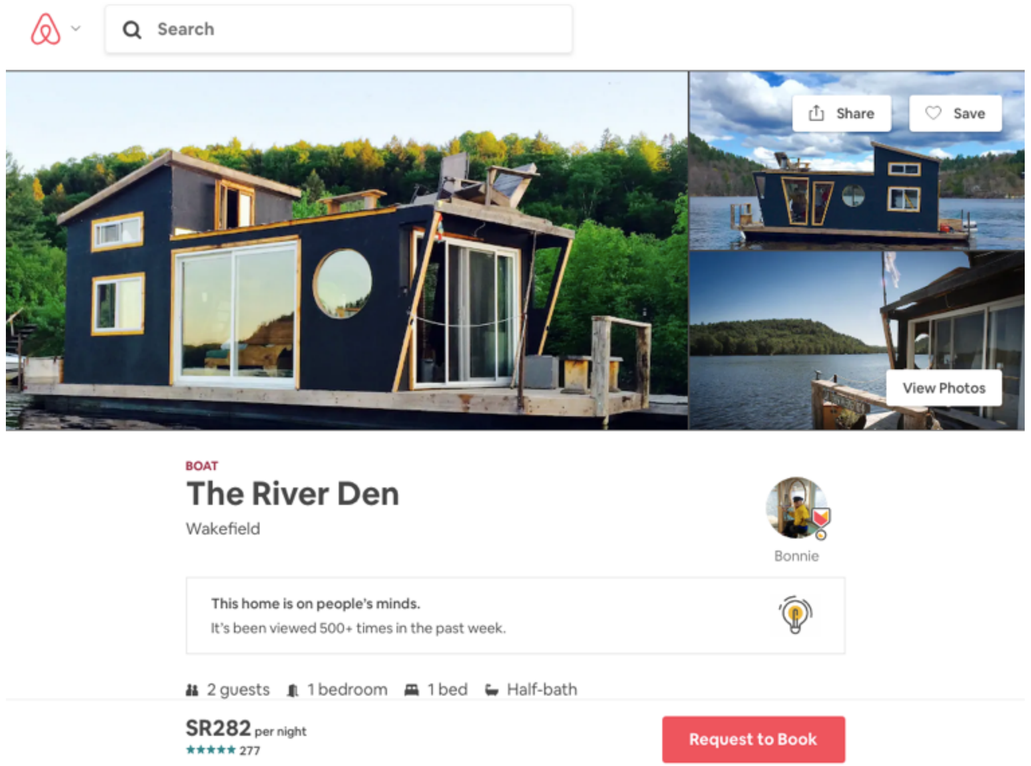


homepage

`listing #104`

Site Search

`listing #73`



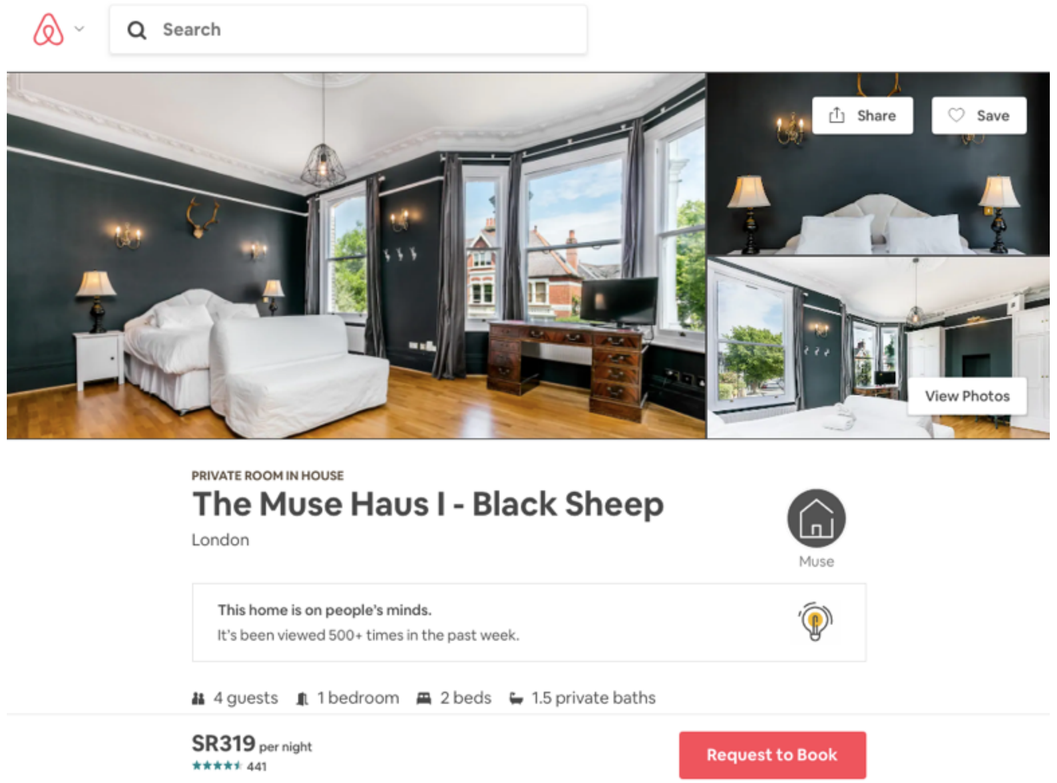
homepage

`listing #104`

Site Search

`listing #73`

`listing #3065`



homepage

`listing #104`

Site Search

`listing #73`

`listing #3065`

`listing #104`

`listing #73`

`listing #3065`

Click sessions

#104	#73	#3065

#100	#200	#300	#400	#500	#600	#700

#100	#200	#300	#400	#500	#600	#700	#800	#900	#1000	#1100	#1200

#100	#200	#300	#200	#500	#600	#700	...

Input listing	output listing	class
#200	#100	1
#200	#300	1

#100	#200	#300	#400	#500	#600	#700	...

Input listing	output listing	class
#200	#100	1
#200	#300	1
#300	#200	1
#300	#400	1
...	...	...

#100	#200	#300	#400	#500	#600	#700	...

Input listing	output listing	class
#200	#100	1
#200	#417	0
#200	#300	1
#200	#71	0
#300	#200	1
#300	#33	0
#300	#400	1
#200	#22	0

Listing #1

Listing #2

Listing #3

Listing #4

Listing #5

Listing #6

Listing #7

…

Listing #4000

listing #3

Listing #1

Listing #2

Listing #3

Listing #4

Listing #5

Listing #6

Listing #7

…

Listing #4000

listing #3

Listing #3

Most similar listings:

Listing #72

Listing #2006

Listing #1345

Listing #491

Listing #304

Listing #1

Listing #2

Listing #3

Listing #4

Listing #5

Listing #6

Listing #7

…

Listing #4000

`listing #3`

Similar listings:

ENTIRE CHALET · MONTE VERDE

Charming place in the Mountain!

SR485 per night

★★★★★ 109

ENTIRE COTTAGE · SETOR SOCIOECONÔMICO 21

Refúgio nas nuvens

SR400 per night

★★★★★ 65

PRIVATE ROOM · SÃO JOSÉ DOS CAMPOS

Saint Barbara High SFX Chalet 1

SR745 per night

★★★★★ 53

Improving Recommendations

listing #3

Listing #3

Most similar listings:

Listing #72

Listing #2006

Listing #1345

listing #3

Listing #3

Most similar listings:

Listing #72

Listing #2006

Listing #1345

After recommending:

clicked

clicked

not clicked

listing #3

Listing #3

Most similar listings:

Listing #72

Listing #2006

Listing #1345

After recommending:

clicked

clicked

not clicked

Input listing	output listing	class
#3	**#1345**	**0**

Conversion as Global Context

Click sessions

#104	#73	#3065

#100	#200	#300	#400	#500	#600	#700

#100	#200	#300	#400	#500	#600	#700	#800	#900	#1000	#1100	#1200

Booked

#100	#200	#300	#200	#500	#600	…	#1200

Input listing	output listing	class
#200	#100	1
#200	#300	1
#200	#1200	1

Booked

#100	#200	#300	#400	#500	#600	…	#1200

Input listing	output listing	class
#200	#100	1
#200	#300	1
#200	#1200	1
#300	#200	1
#300	#400	1
#300	#1200	1

Real-time Personalization using Embeddings for Search Ranking at Airbnb

Mihajlo Grbovic
Airbnb, Inc.
San Francisco, California, USA
mihajlo.grbovic@airbnb.com

Haibin Cheng
Airbnb, Inc.
San Francisco, California, USA
haibin.cheng@airbnb.com

https://www.kdd.org/kdd2018/accepted-papers/view/real-time-personalization-using-embeddings-for-search-ranking-at-airbnb

Embedding Evaluation Tool

Search

Query Type

Listing ID

Listing ID

16486364

Search

I'm Feeling Lucky

$236 Cabane Secrète pour 2 personnes

Location: Raray, Hauts-de-France, France
Description: Mon logement est proche de la nature et du golf. Vous apprécierez mon logement pour sa tranquillité et son confort. Mon logement est parfait pour les ...

Embedding Evaluation Tool

Search

Query Type

Listing ID

Listing ID

16486364

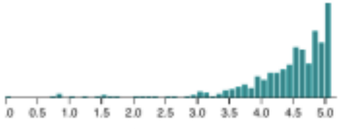
Search

I'm Feeling Lucky

$236 Cabane Secrète pour 2 personnes

Location: Raray, Hauts-de-France, France
Description: Mon logement est proche de la nature et du golf. Vous apprécierez mon logement pour sa tranquillité et son confort. Mon logement est parfait pour les ...

Score Histogram



*Includes scores for up to the 500 nearest listings

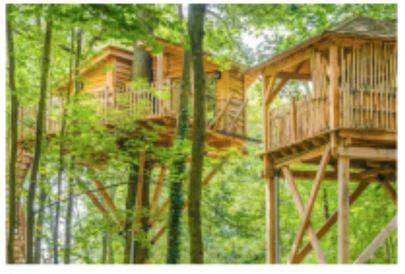
Nearest listings (10)

$236 Cabane Secrète pour 2 personnes

KNN: /admin/embedding_evaluation/16486364
Score: 0.00
Location: Raray, Hauts-de-France, France
Description: Mon logement est proche de la nature et du golf. Vous apprécierez mon logement pour sa tranquillité et son confort. Mon logement est parfait pour les ...

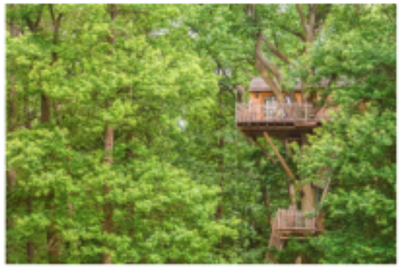
$386 Cabane Spa Origin

KNN: /admin/embedding_evaluation/16905264
Score: 0.70
Location: Raray, Hauts-de-France, France
Description: Mon logement est proche de la nature et de la forêt. Vous apprécierez mon logement pour l'emplacement, les espaces extérieurs et sa tranquillité. Mon ...



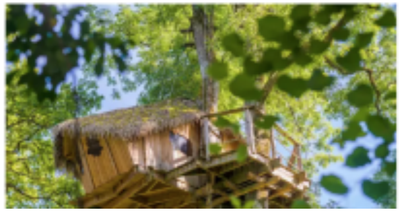
$320 Cabane SPA Cocon pour 2 personnes

KNN: /admin/embedding_evaluation/16486854
Score: 0.84
Location: Raray, Hauts-de-France, France
Description: Mon logement est proche de Paris . Vous apprécierez mon logement pour son bain nordique privé. Mon logement est parfait pour les couples.

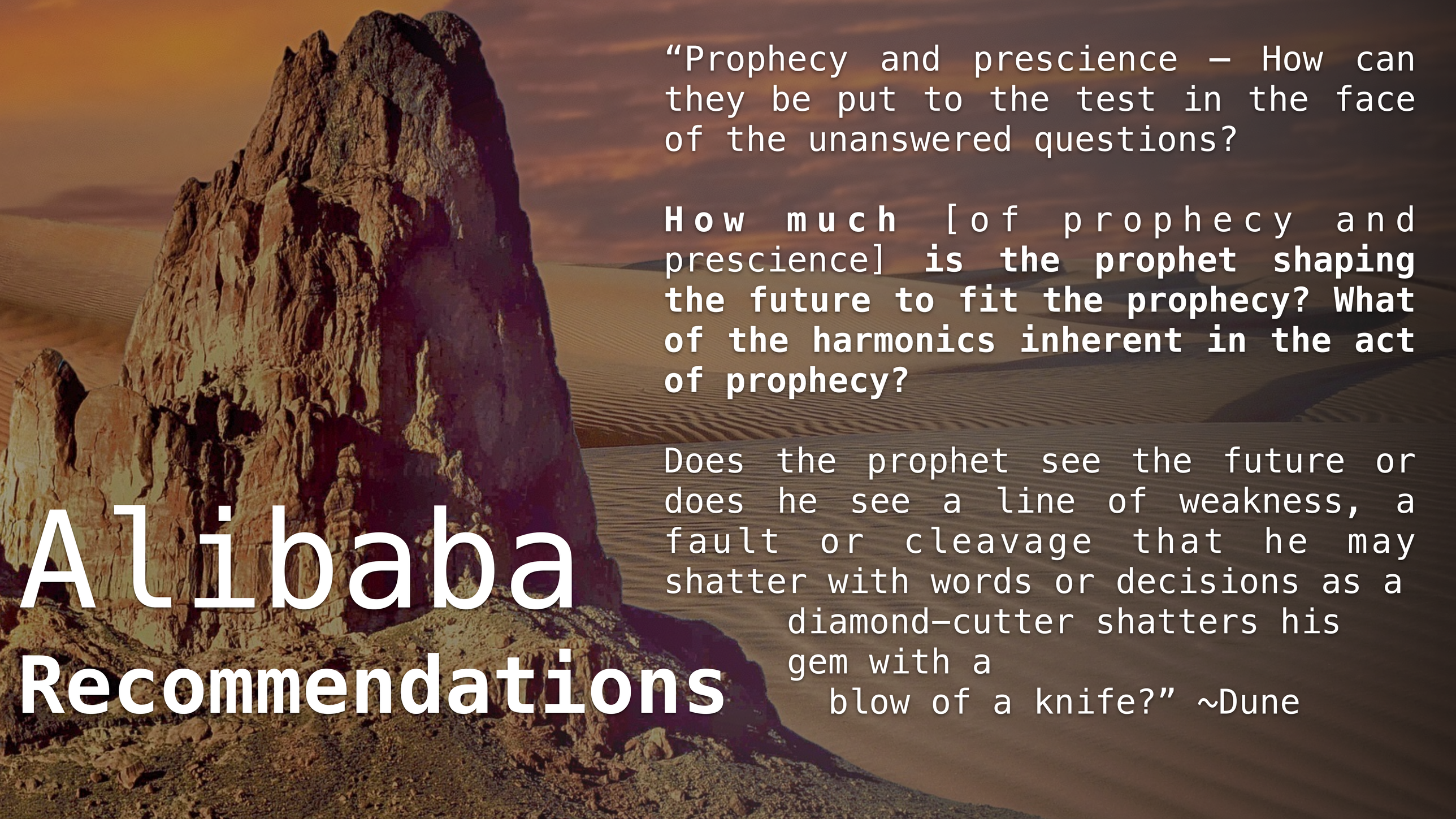


$236 Cabane Imprenable pour 2 personnes

KNN: /admin/embedding_evaluation/16485735
Score: 0.87
Location: Raray, Hauts-de-France, France
Description: Mon logement est proche du golf et du château. Vous apprécierez mon logement pour son calme et son confort. Mon logement est parfait pour les couples....



The A/B test showed that embedding-based solution lead to a 21% increase in Similar Listing carousel CTR (23% in cases when listing page had entered dates and 20% in cases of dateless pages) and 4.9% increase in guests who find the listing they end up booking in Similar Listing carousel. In light of these results we deployed the embedding-based Similar Listings to production.



Alibaba Recommendations

“Prophecy and prescience — How can they be put to the test in the face of the unanswered questions?

How much [of prophecy and prescience] **is the prophet shaping the future to fit the prophecy? What of the harmonics inherent in the act of prophecy?**

Does the prophet see the future or does he see a line of weakness, a fault or cleavage that he may shatter with words or decisions as a diamond-cutter shatters his gem with a blow of a knife?” ~Dune

#100	#200	#500

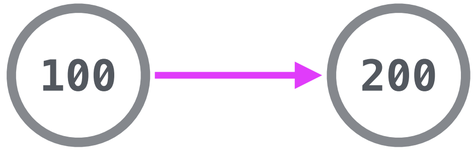
#100	#400	#100	#400	#500

#100	#300	#500	#200

#100	#200	#500

#100	#400	#100	#400	#500

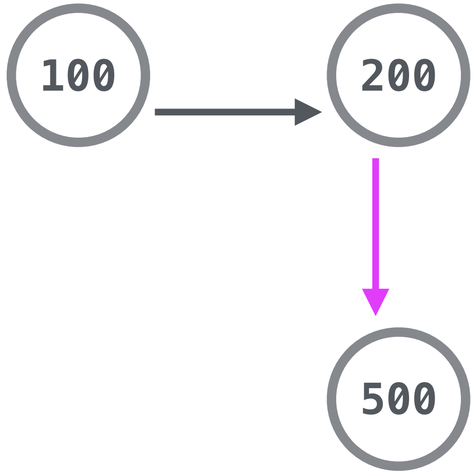
#100	#300	#500	#200



#100	#200	#500

#100	#400	#100	#400	#500

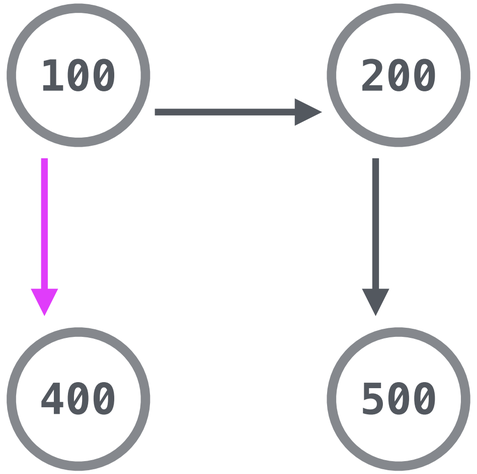
#100	#300	#500	#200



#100	#200	#500

#100	#400	#100	#400	#500

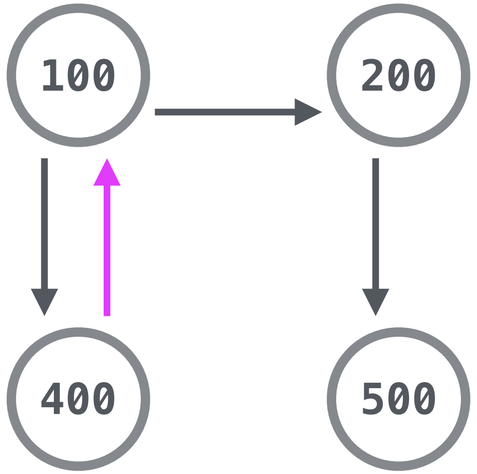
#100	#300	#500	#200



#100	#200	#500

#100	#400	#100	#400	#500

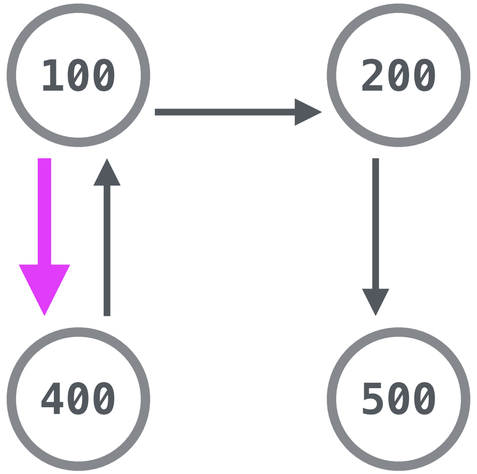
#100	#300	#500	#200



#100	#200	#500

#100	#400	#100	#400	#500

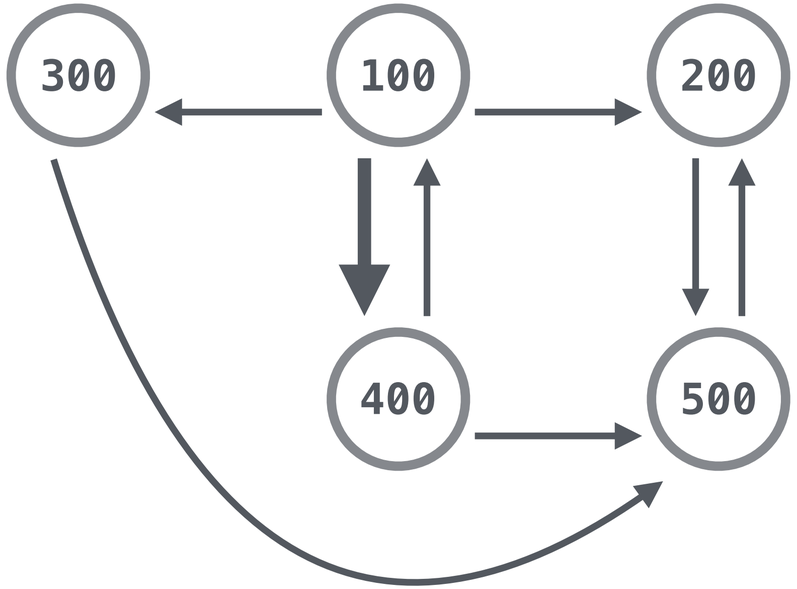
#100	#300	#500	#200

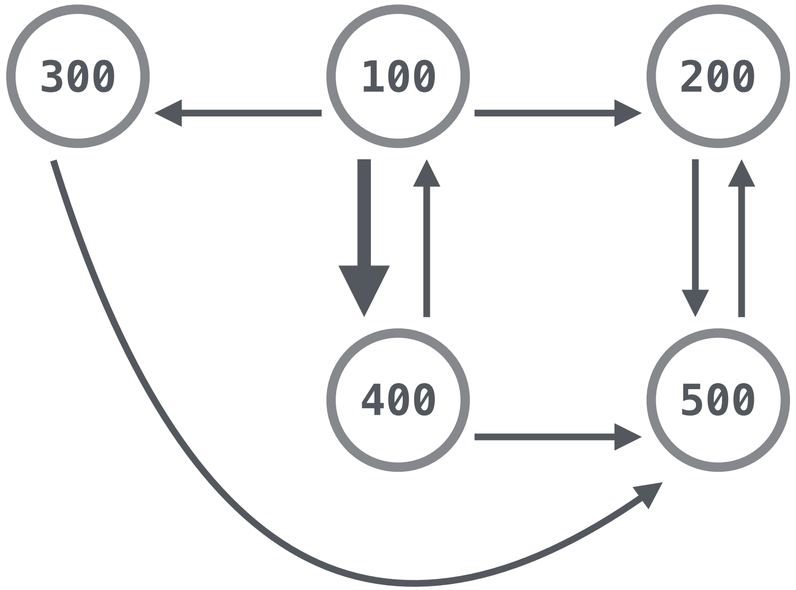


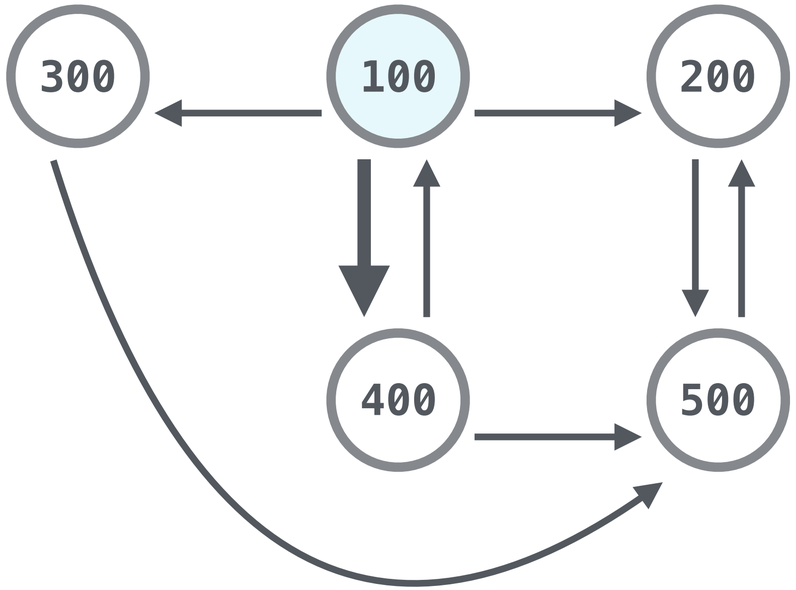
#100	#200	#500

#100	#400	#100	#400	#500

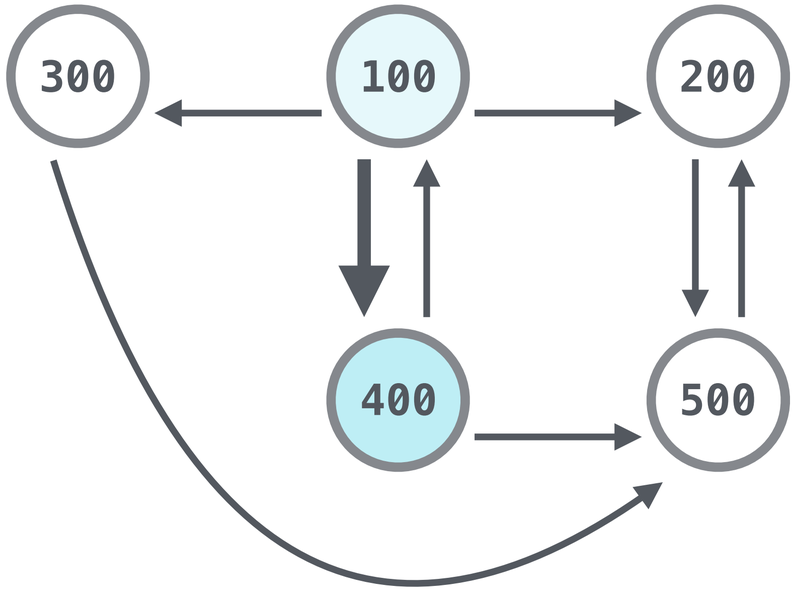
#100	#300	#500	#200



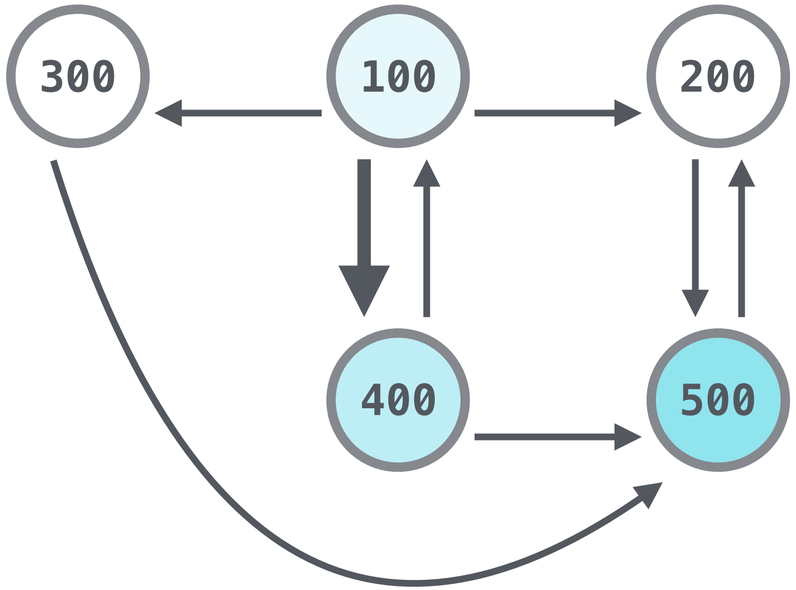

300
100
200
400
500



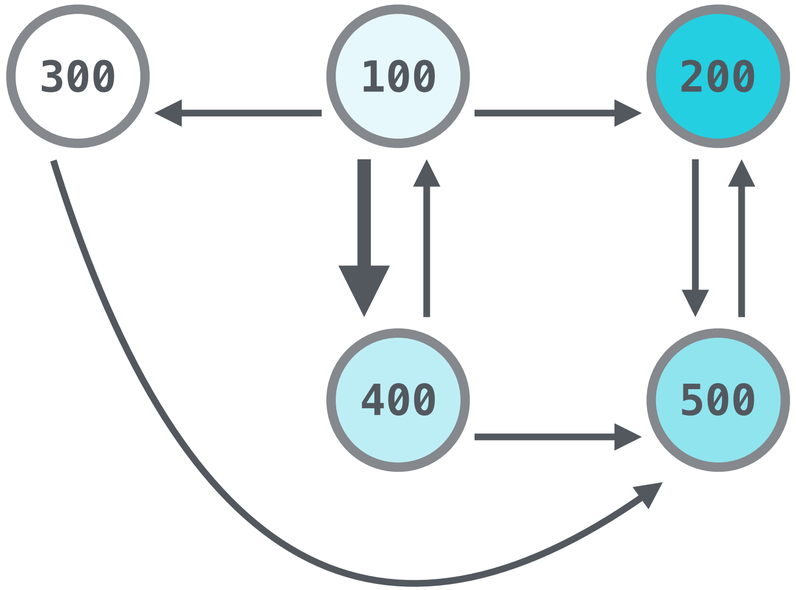
#100



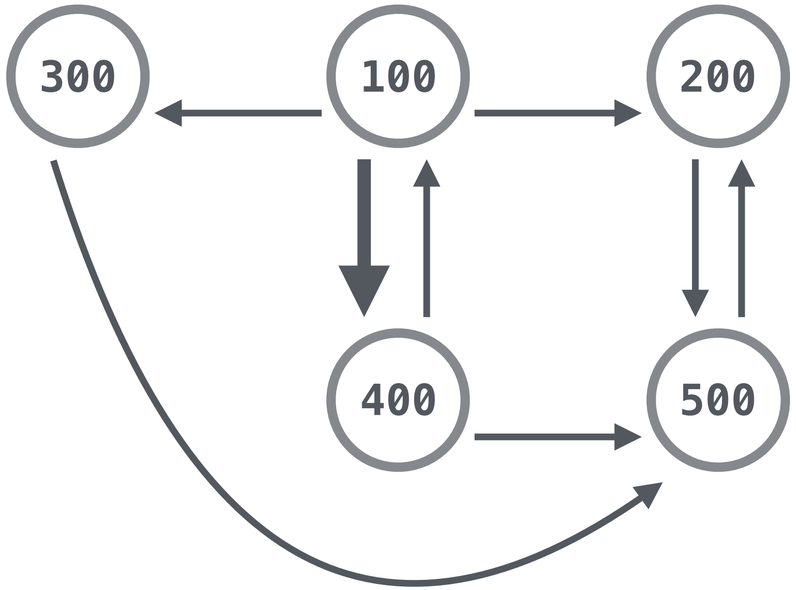
#100	#400



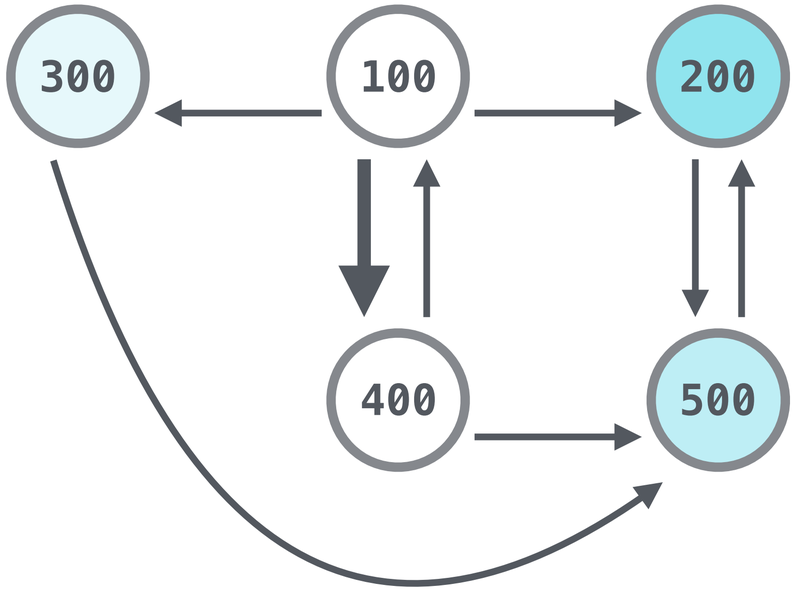
#100	#400	#500



#100	#400	#500	#200

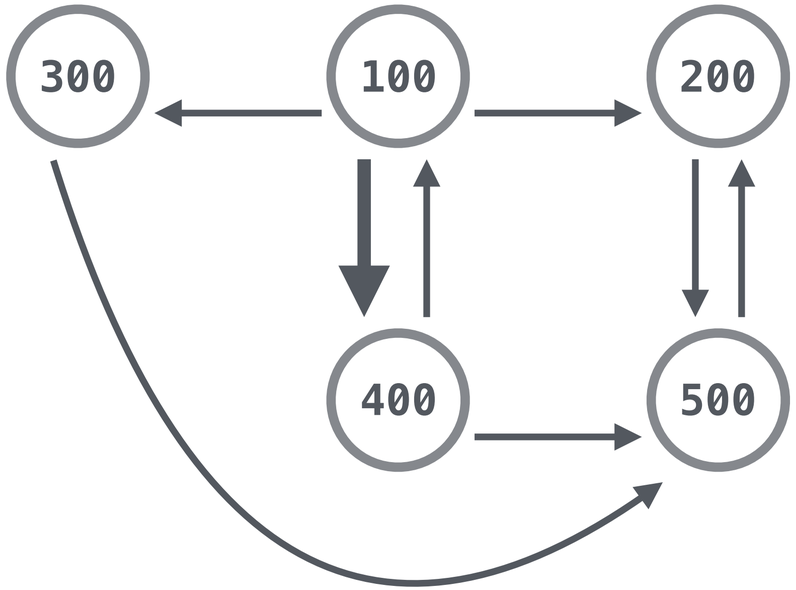


#100	#400	#500	#200



#100	#400	#500	#200

#300	#500	#200



#100	#400	#500	#200

#300	#500	#200

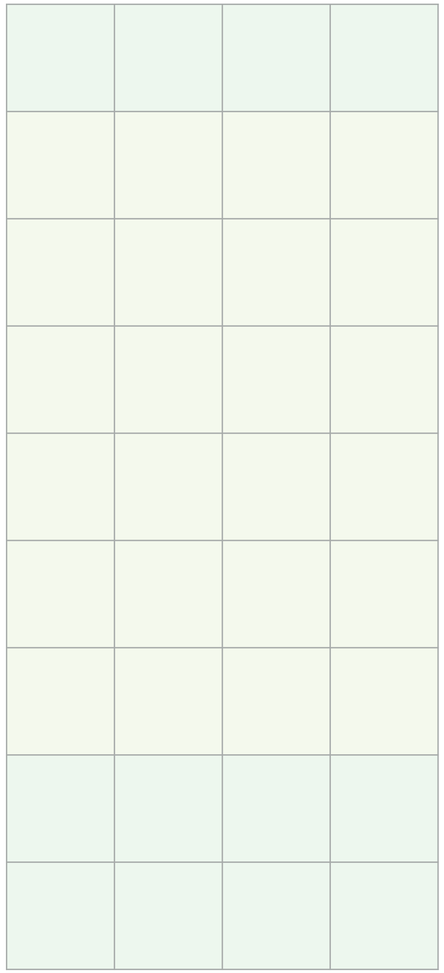
#400	#100	#300	#500	#200	#500

#100	#400	#500	#200

Input item	output item	class
#400	#100	1
#400	#500	1

#100	#400	#500	#200

Input item	output item	class
#400	#100	1
#400	#500	1
#500	#400	1
#500	#200	1



Item #1

Item #2

Item #3

Item #4

Item #5

Item #6

Item #7

…

Item #4000

Billion-scale Commodity Embedding for E-commerce Recommendation in Alibaba

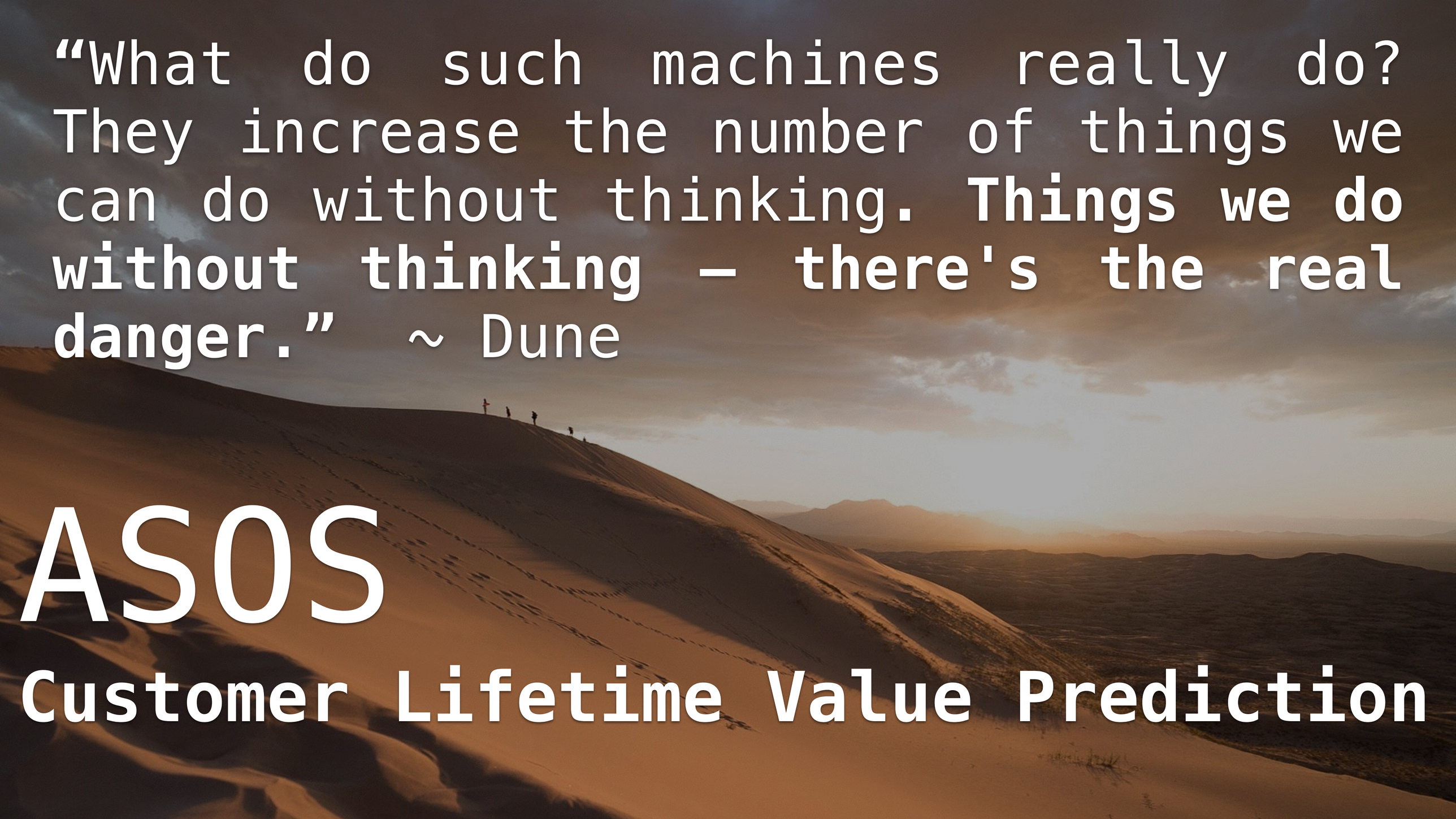
Jizhe Wang, Pipei Huang*
Alibaba Group
Hangzhou and Beijing, China
{jizhe.wjz,pipei.hpp}@alibaba-inc.com

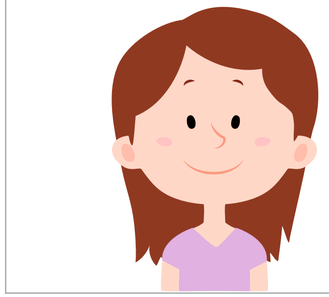
Huan Zhao
Department of Computer Science and Engineering
Hong Kong University of Science and Technology
Kowloon, Hong Kong
hzhaoaf@cse.ust.hk

Zhibo Zhang, Binqiang Zhao
Alibaba Group
Beijing, China
{shaobo.zzb,binqiang.zhao}@alibaba-inc.com

Dik Lun Lee
Department of Computer Science and Engineering
Hong Kong University of Science and Technology
Kowloon, Hong Kong
dlee@cse.ust.hk

https://www.kdd.org/kdd2018/accepted-papers/view/billion-scale-commodity-embedding-for-e-commerce-recommendation-in-alibaba

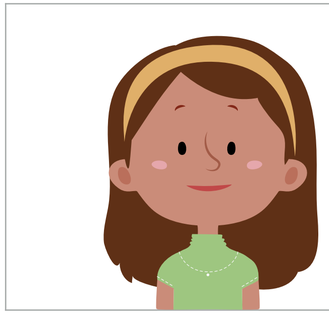

“What do such machines really do? They increase the number of things we can do without thinking. **Things we do without thinking — there's the real danger.”** ~ Dune
ASOS
Customer Lifetime Value Prediction



User #200

User #300

User #500

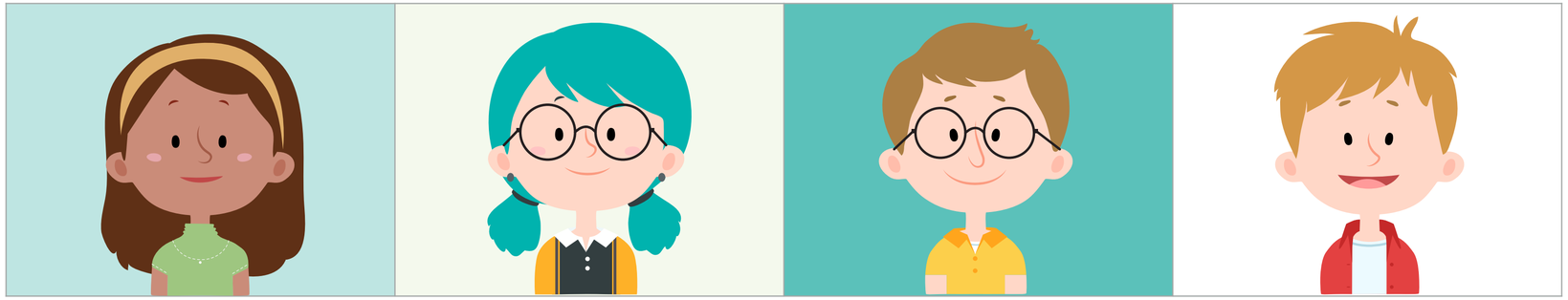


User #300

User #700

User #400

User #100



User #300 User #700 User #400 User #100

Input user	output user	class
#700	#300	1
#700	#400	1

Customer Lifetime Value Prediction Using Embeddings

Benjamin Paul Chamberlain
Department of Computing
Imperial College London
b.chamberlain14@imperial.ac.uk

Ângelo Cardoso
ASOS.com
London, United Kingdom

C.H. Bryan Liu
ASOS.com
London, United Kingdom

Roberto Pagliari
ASOS.com
London, United Kingdom

Marc Peter Deisenroth
Department of Computing
Imperial College London
m.deisenroth@imperial.ac.uk

https://www.kdd.org/kdd2018/accepted-papers/view/perceive-your-users-in-depth-learning-universal-user-representations-from-m

Other Examples:

Anghami - Using Word2vec for Music Recommendations
https://towardsdatascience.com/using-word2vec-for-music-recommendations-bb9649ac2484

Spotify - Machine learning @ Spotify
https://www.slideshare.net/AndySloane/machine-learning-spotify-madison-big-data-meetup

Resources:

Speech and Language Processing [Book]
Dan Jurafsky and James H. Martin
https://web.stanford.edu/~jurafsky/slp3/

Neural Network Methods for Natural Language Processing [Book]
Yoav Goldberg
https://www.amazon.com/dp/B071FGKZMH/

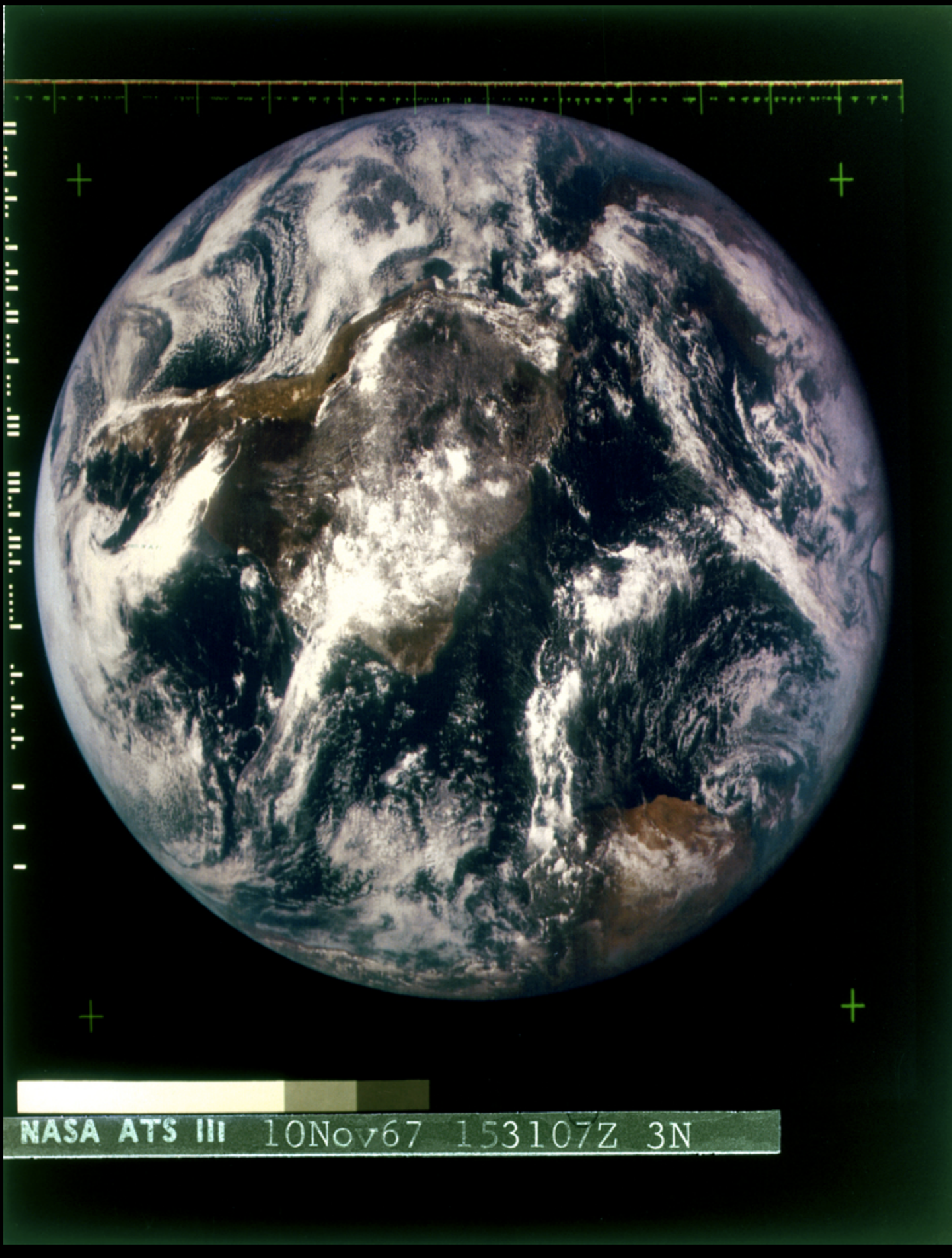
Chris McCormick [Blog]
http://mccormickml.com/

“The thing the ecologically illiterate don’t realize about an ecosystem is that it’s a system. A system! **A system maintains a certain fluid stability that can be destroyed by a misstep in just one niche.**

A system has order, a flowing from point to point. If something dams the flow, order collapses. The untrained might miss that collapse until it was too late. **That’s why the highest function of ecology is the understanding of consequences.**” ~Dune

"first planetary ecology novel
 on a grand scale”
The Cambridge Companion to Science Fiction

"Environmentalists have pointed out that Dune's popularity as a novel depicting a planet as a complex—almost living—thing, in **combination with the first images of Earth from space being published in the same time period,** strongly influenced environmental movements such as the establishment of the international Earth Day.”
Wikipedia


NASA ATS III 10Nov67 153107Z 3N

Justin Shanes
@justinshanes

Follow

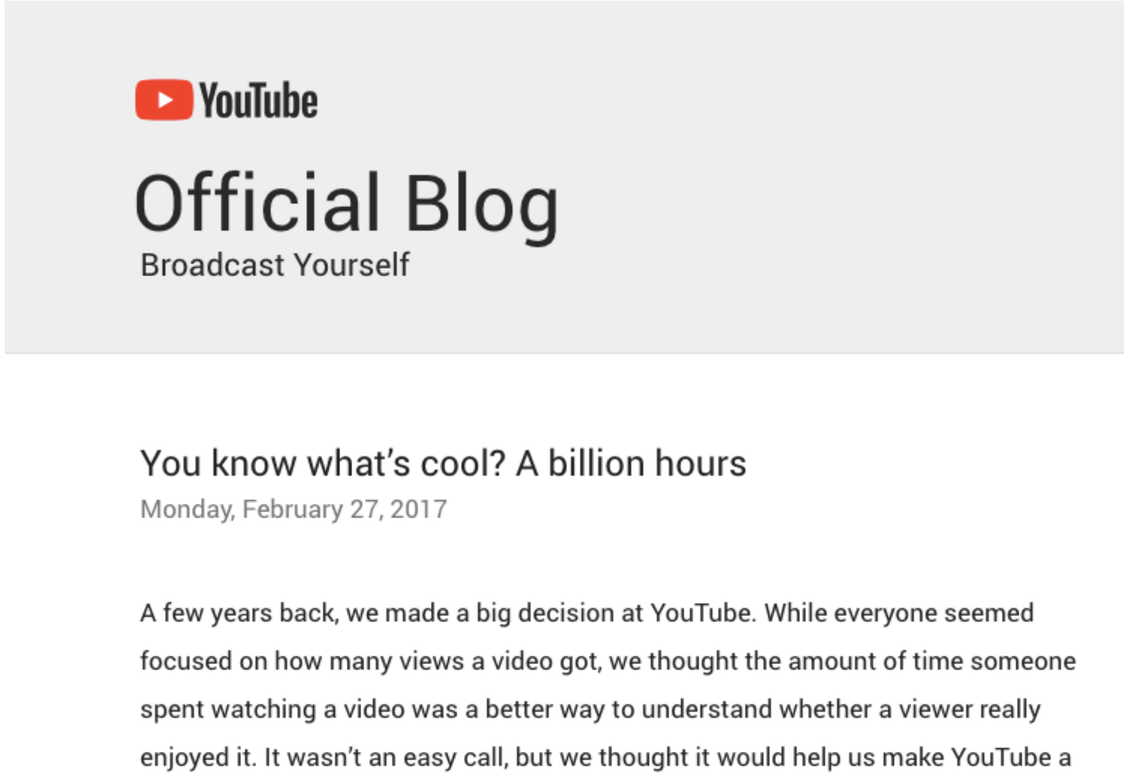
Amazon thinks my recent humidifier purchase was merely the inaugural move in a newfound hobby of humidifier collecting.

7:18 AM - 29 Nov 2016

10,182 Retweets **29,093** Likes

235 10K 29K

People watch *one billion hours* of YouTube every *day*

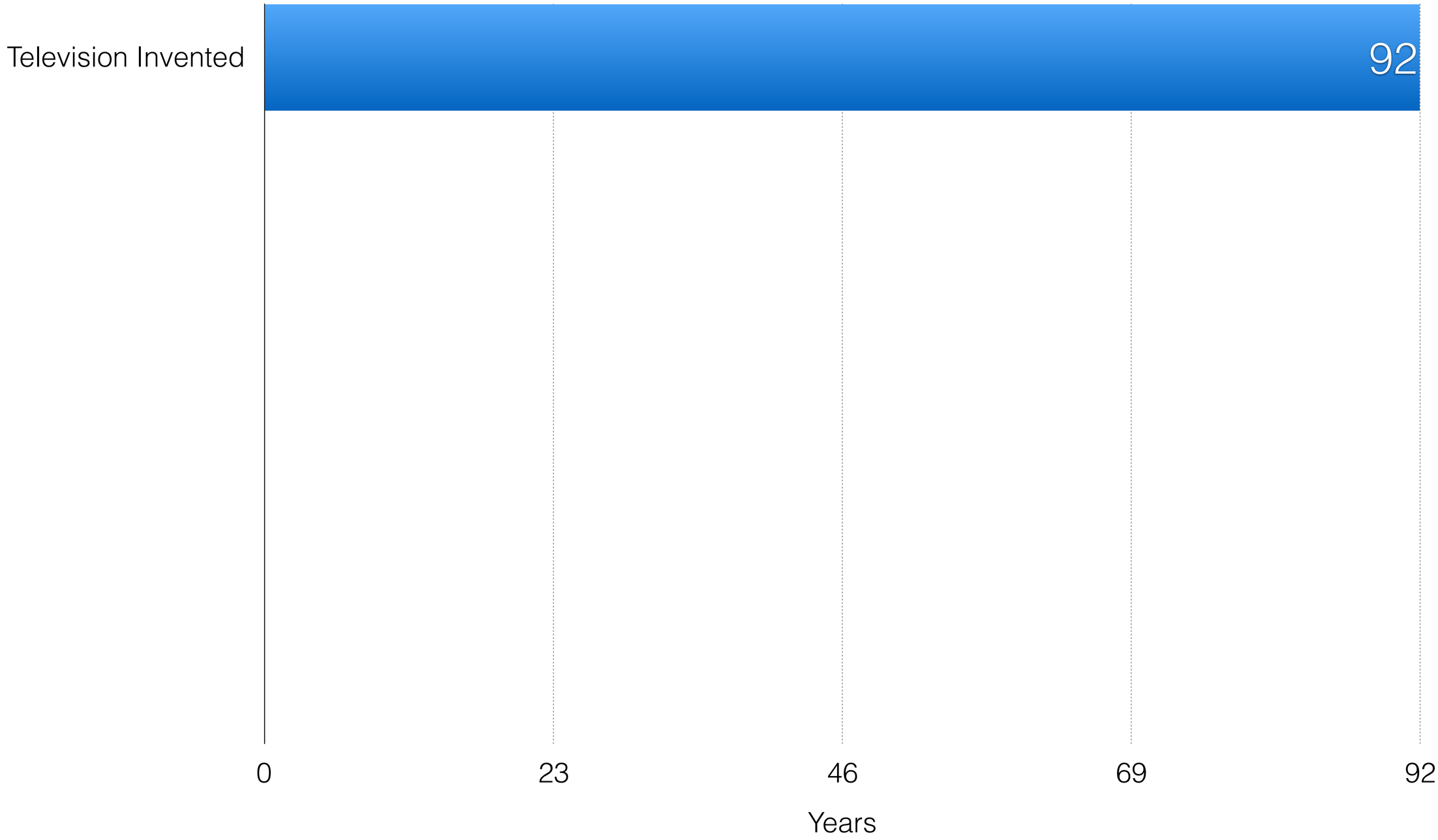


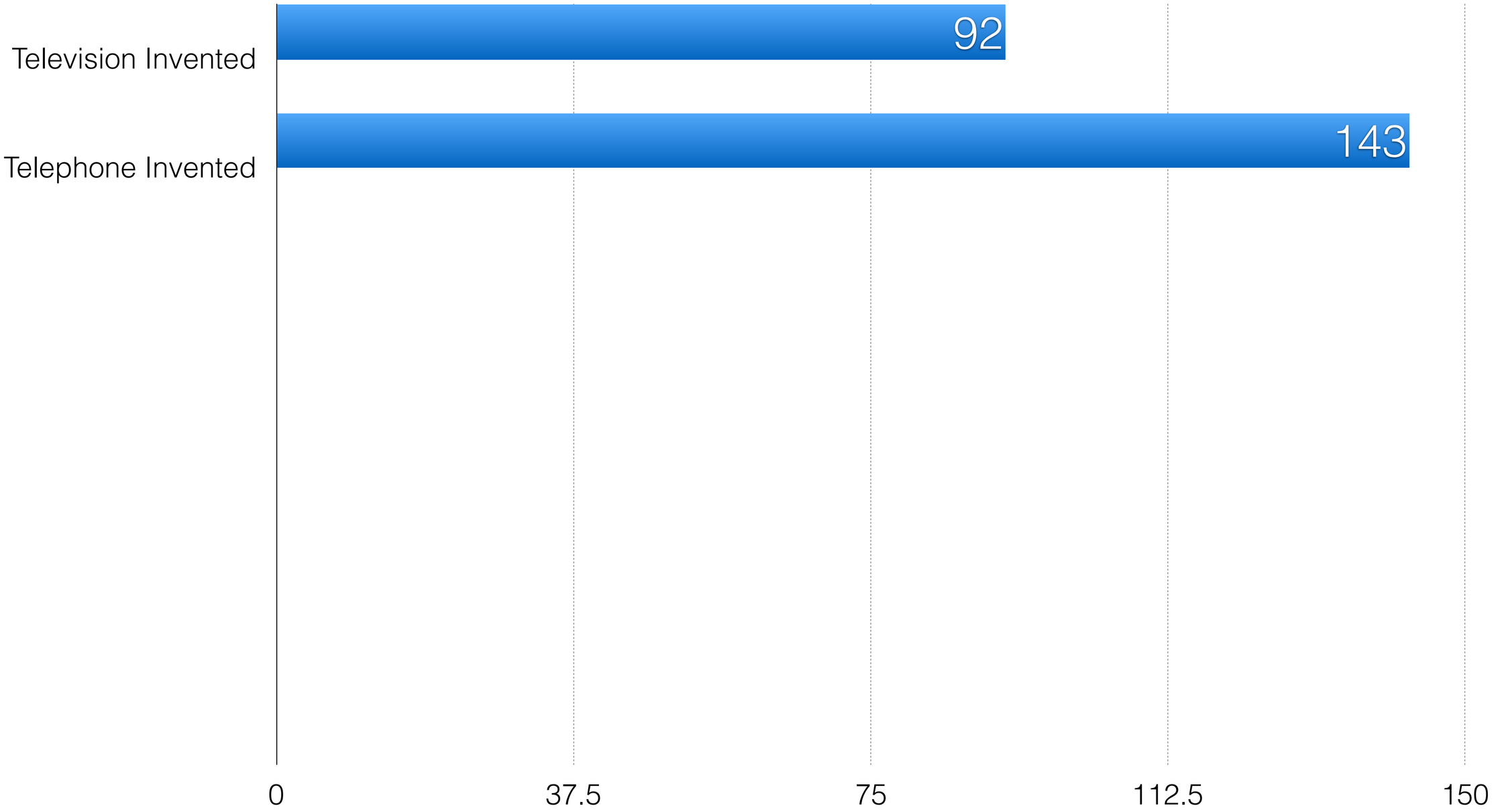
https://youtube.googleblog.com/2017/02/you-know-whats-cool-billion-hours.html

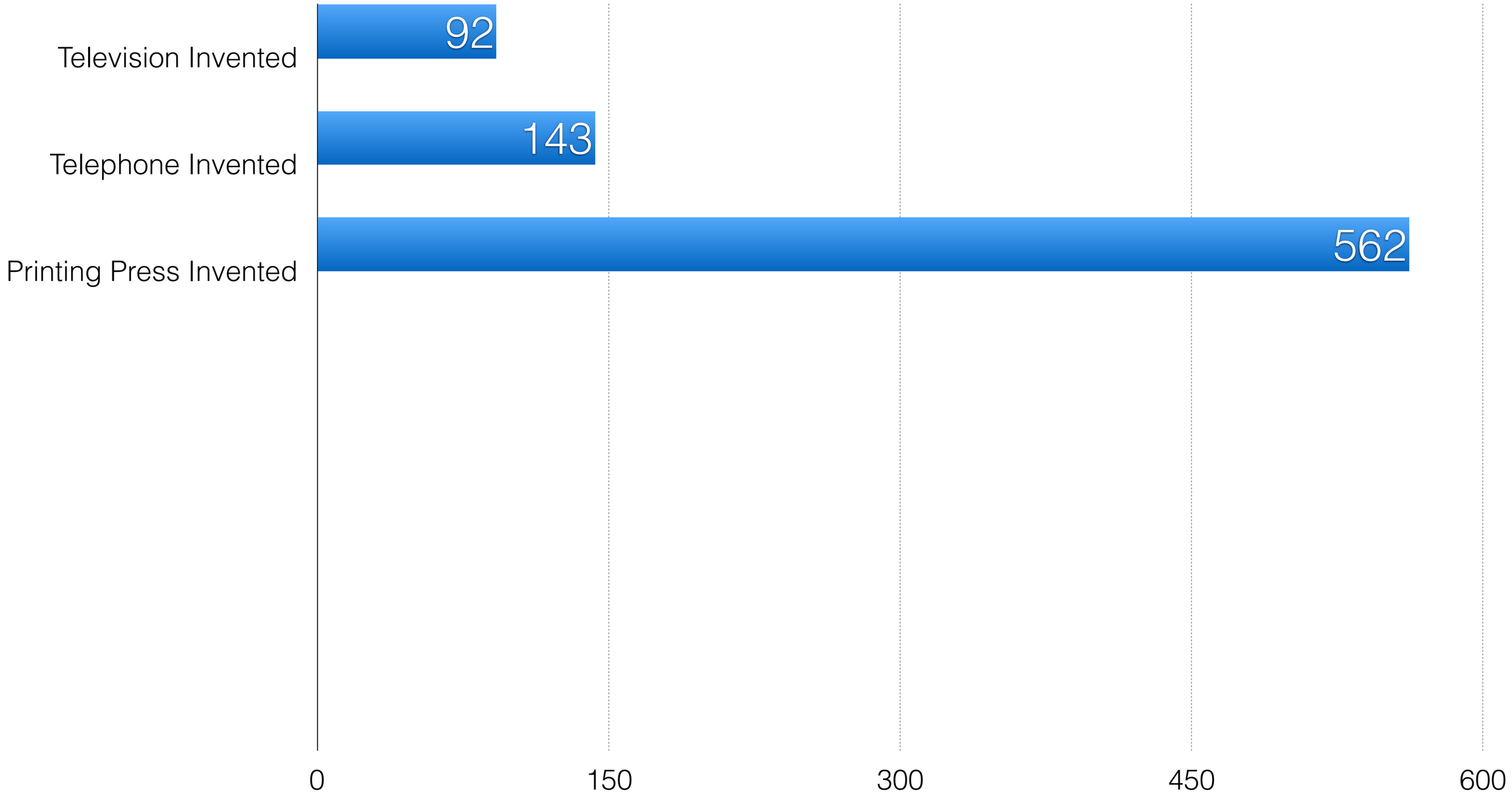
https://qz.com/1178125/youtubes-recommendations-drive-70-of-what-we-watch/

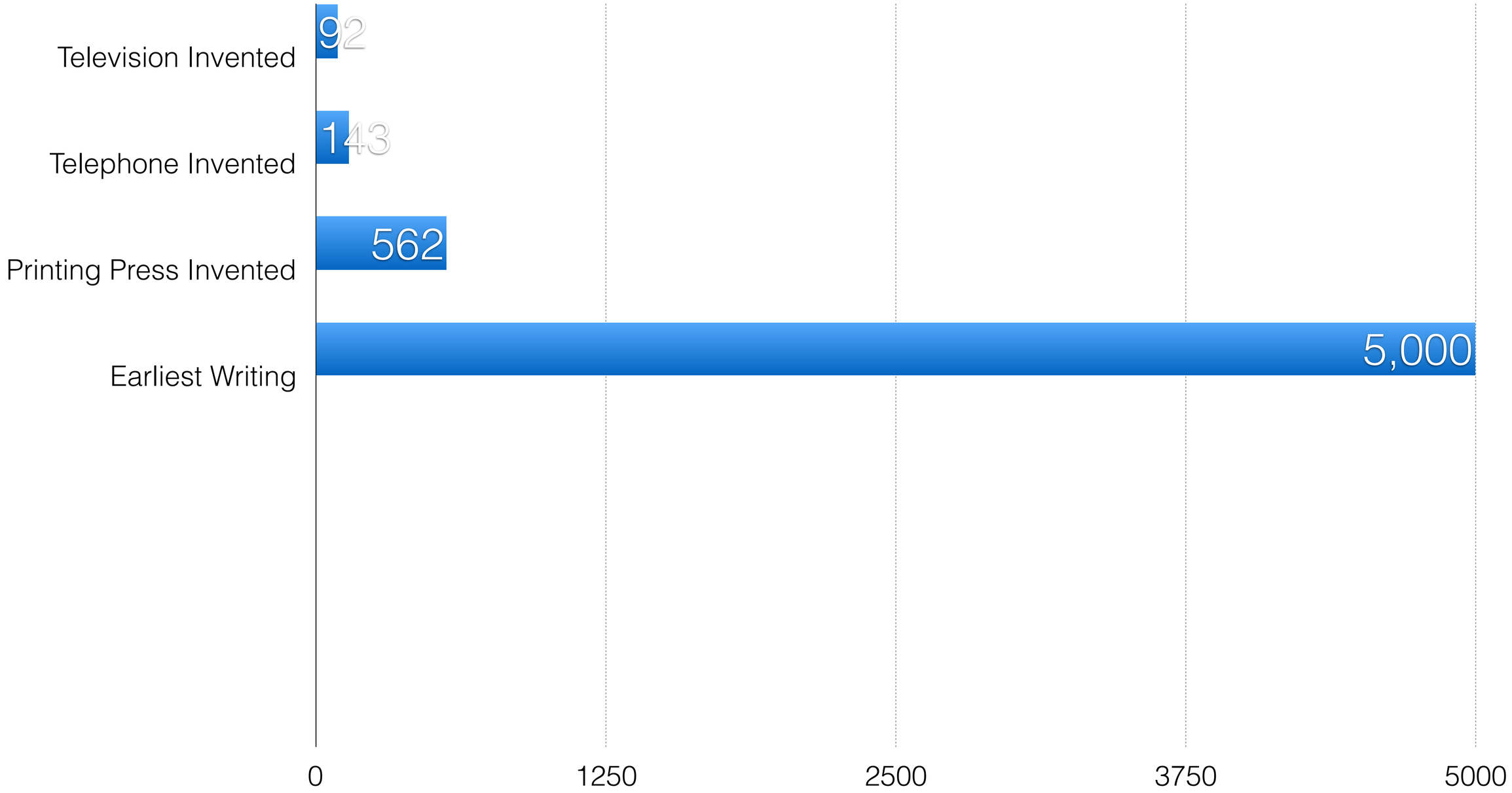
700,000,000 hours of video / day

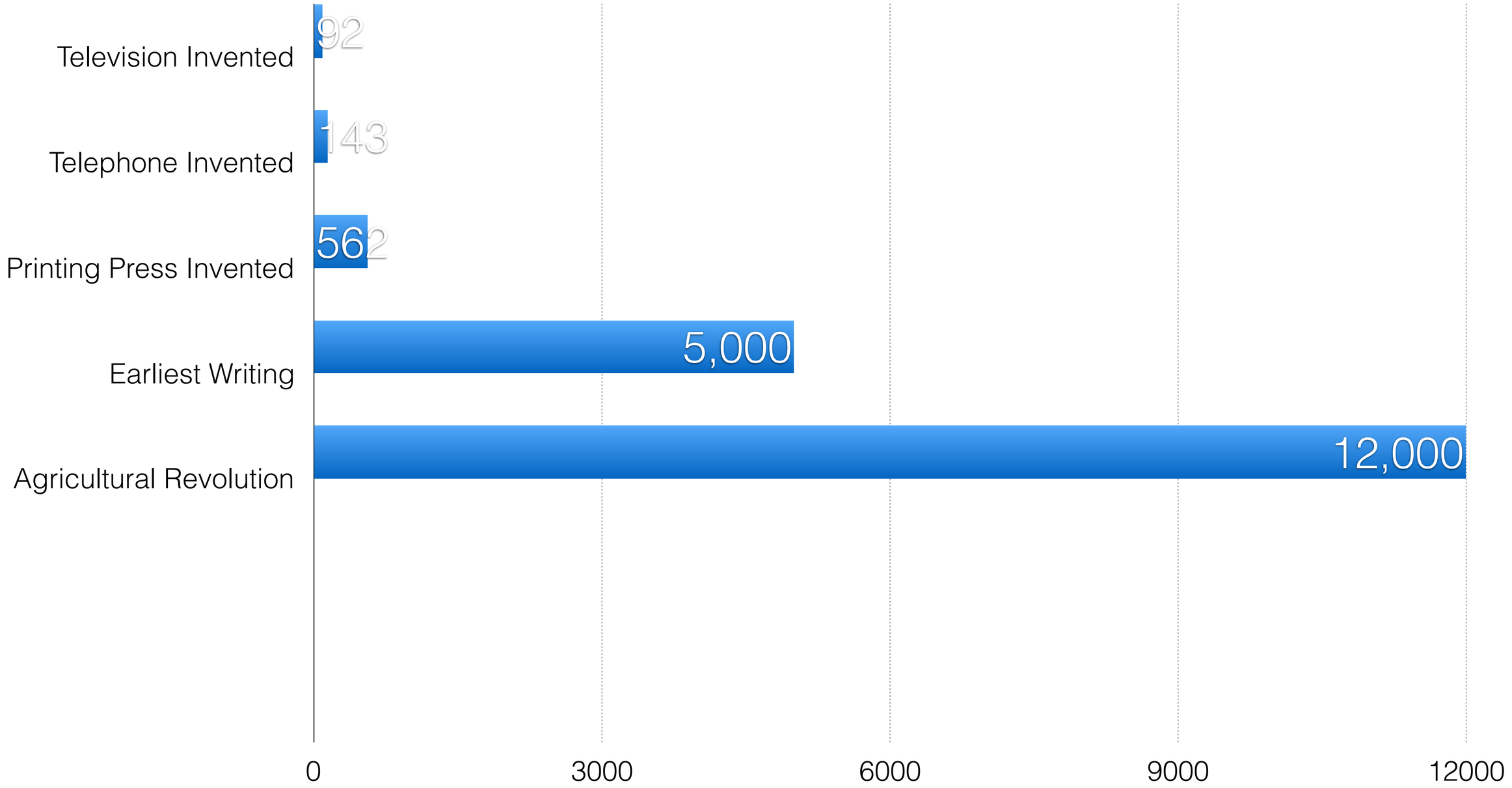
https://algotransparency.org

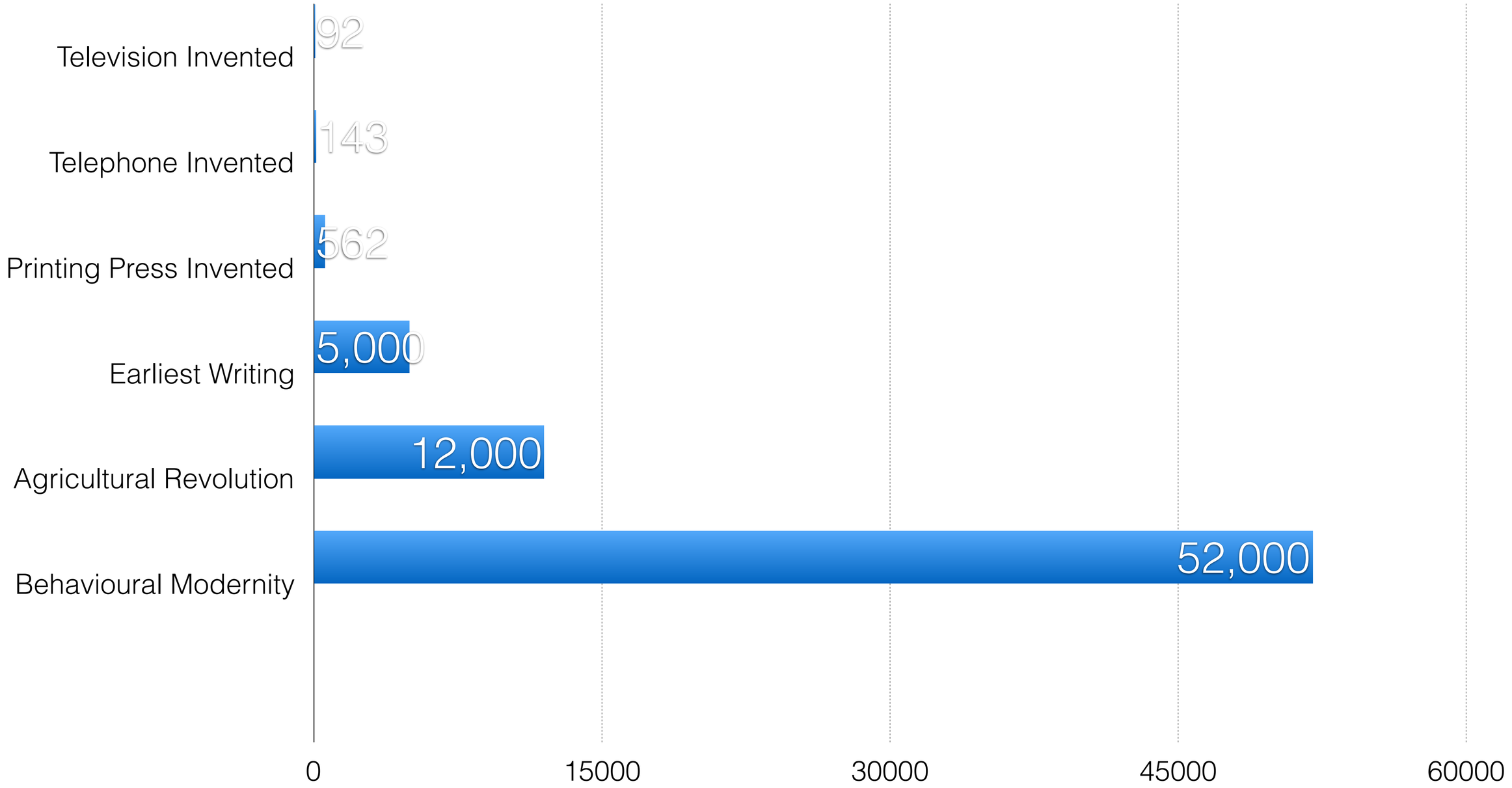

Television Invented
92
0
23
46
69
92
Years

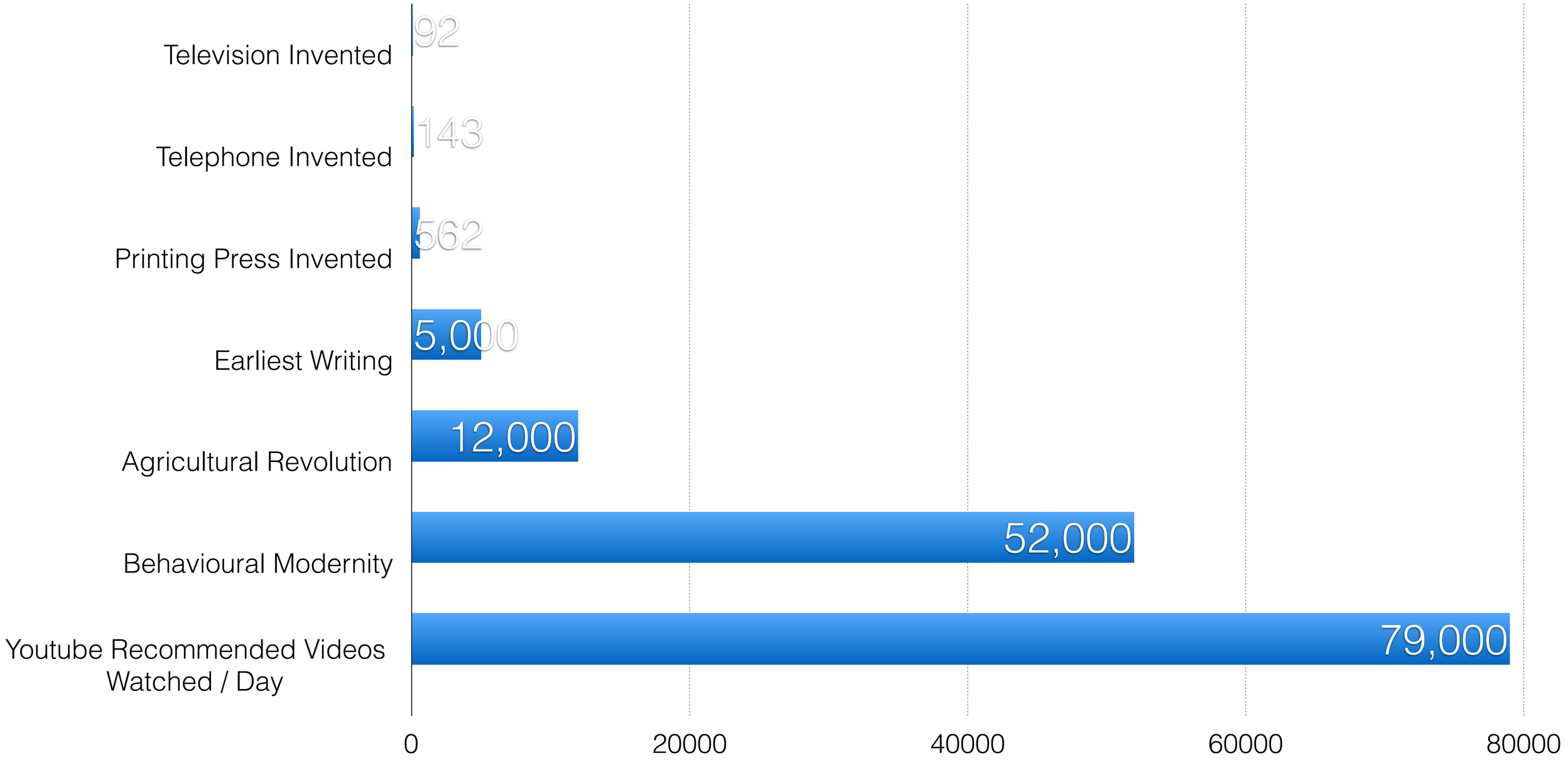

Television Invented
92
Telephone Invented
143
0
37.5
75
112.5
150


Television Invented
92
Telephone Invented
143
Printing Press Invented
562
0
150
300
450
600


Television Invented
92
Telephone Invented
143
Printing Press Invented
562
Earliest Writing
5,000
0
1250
2500
3750
5000


Television Invented
92
Telephone Invented
143
Printing Press Invented
562
Earliest Writing
5,000
Agricultural Revolution
12,000
0
3000
6000
9000
12000


Television Invented
92
Telephone Invented
143
Printing Press Invented
562
Earliest Writing
5,000
Agricultural Revolution
12,000
Behavioural Modernity
52,000
0
15000
30000
45000
60000


Television Invented
92
Telephone Invented
143
Printing Press Invented
562
Earliest Writing
5,000
Agricultural Revolution
12,000
Behavioural Modernity
52,000
Youtube Recommended Videos Watched / Day
79,000
0
20000
40000
60000
80000

TECHNOLOGY NEWS SEPTEMBER 8, 2017 / 7:04 PM / A YEAR AGO

Two-thirds of American adults get news from social media: survey

Angela Moon

2 MIN READ

(Reuters) - About two-thirds of American adults are getting "at least some of their news on social media" with two-in-ten doing so often, according to a Pew Research Center survey this week.

FRANCE 24

LIVE

WHO warns of 'backsliding' as measles cases soar worldwide

Date created : 15/02/2019 - 07:51

https://www.france24.com/en/20190215-health-science-measles-who-warns-backsliding-cases-soar-worldwide-vaccination-children

FRANCE 24

LIVE

WHO warns of 'backsliding' as measles cases soar worldwide

Date created : 15/02/2019 - 07:51

"case numbers worldwide surging around 50 percent last year"

https://www.france24.com/en/20190215-health-science-measles-who-warns-backsliding-cases-soar-worldwide-vaccination-children

FRANCE 24

LIVE

WHO warns of 'backsliding' as measles cases soar worldwide

Date created : 15/02/2019 - 07:51

"case numbers worldwide surging around 50 percent last year"

"Last year, measles caused approximately 136,000 deaths around the world"

https://www.france24.com/en/20190215-health-science-measles-who-warns-backsliding-cases-soar-worldwide-vaccination-children

FRANCE 24

LIVE

WHO warns of 'backsliding' as measles cases soar worldwide

Date created : 15/02/2019 - 07:51

"case numbers worldwide surging around 50 percent last year"

"Last year, measles caused approximately 136,000 deaths around the world"

"experts blame the problem in part on complacency and misinformation about the vaccine […] which have been spread in part on social media"

https://www.france24.com/en/20190215-health-science-measles-who-warns-backsliding-cases-soar-worldwide-vaccination-children

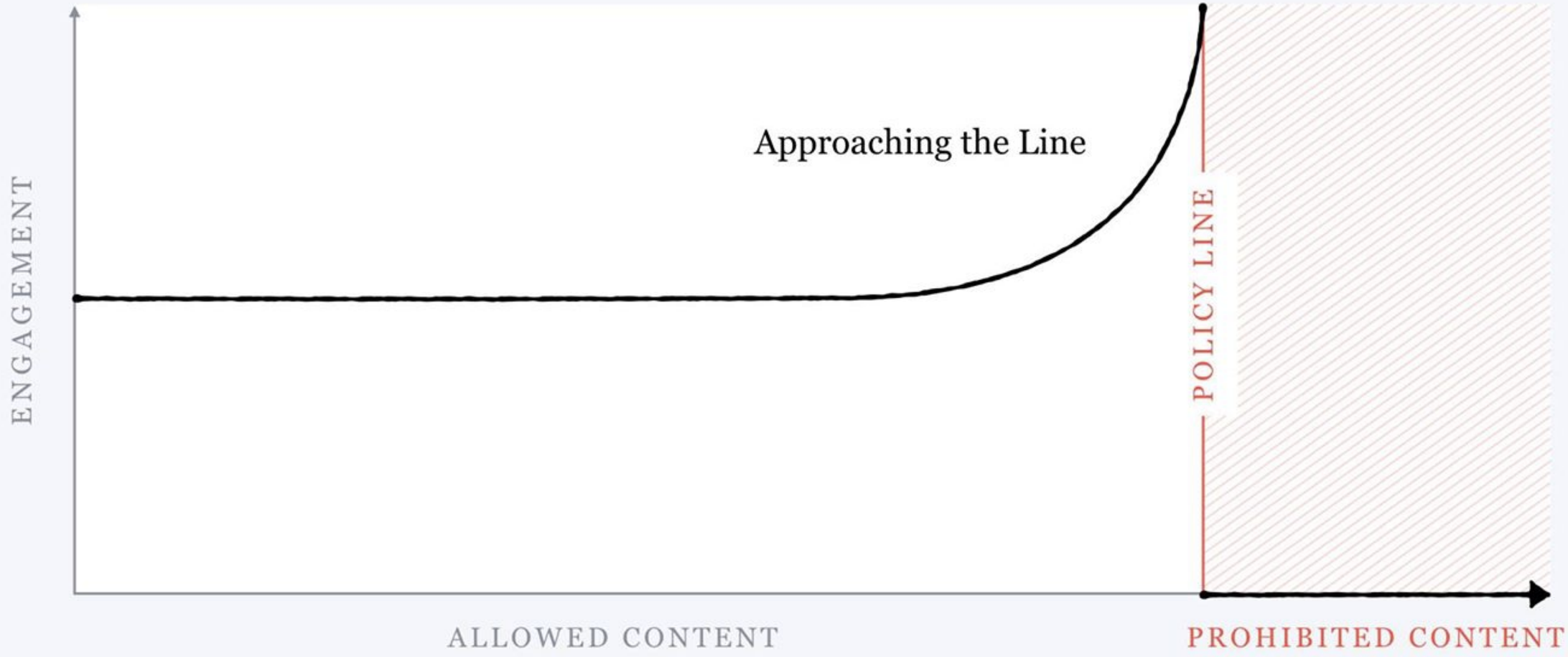
A Blueprint for Content Governance and Enforcement

MARK ZUCKERBERG · THURSDAY, NOVEMBER 15, 2018

My focus in 2018 has been addressing the most important issues facing Facebook. As the year wraps up, I'm writing a series of notes about these challenges and the progress we've made. The first note was about Preparing for Elections and this is the second in the series.

...

Natural Engagement Pattern



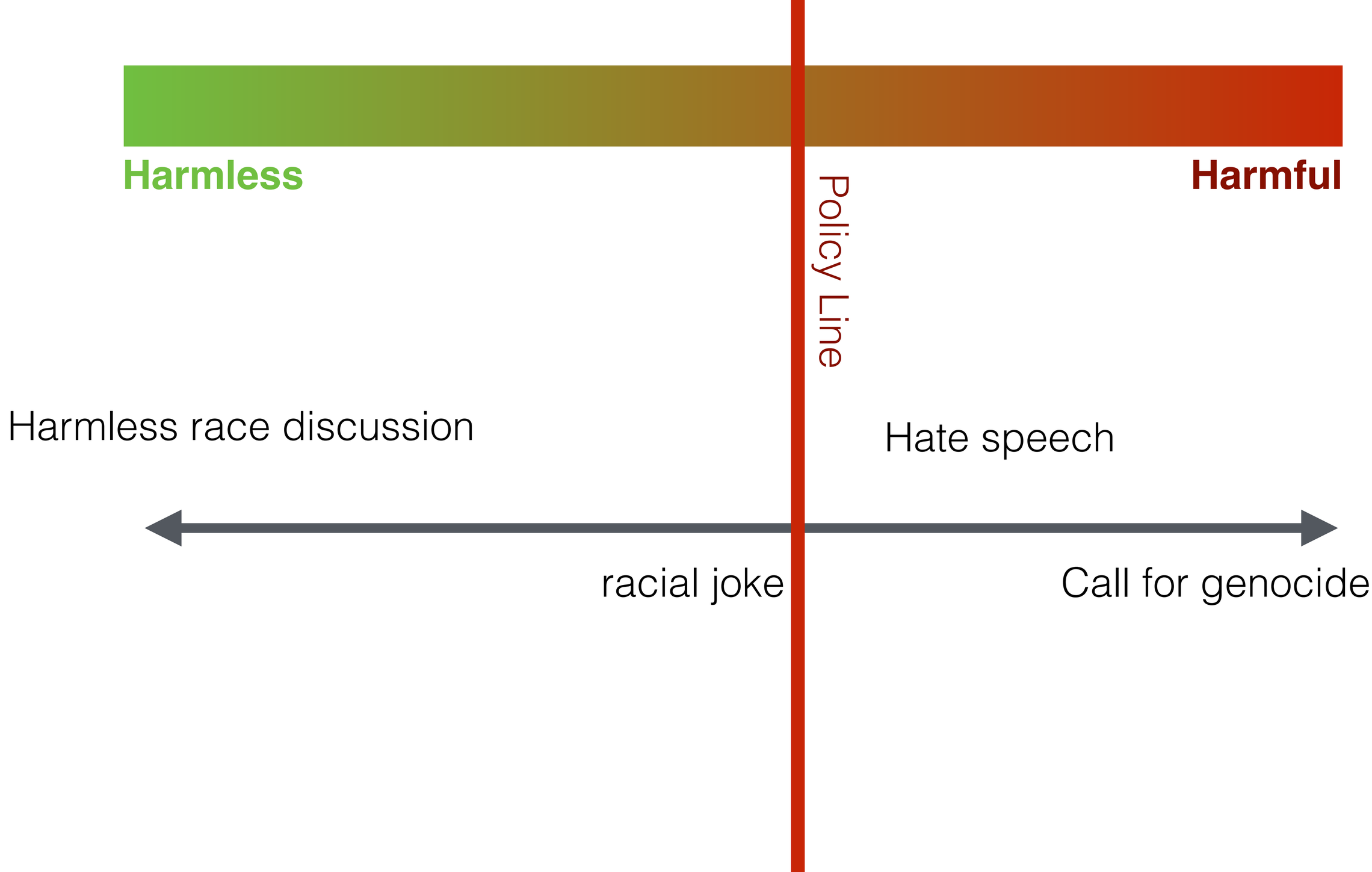
Harmless **Harmful**

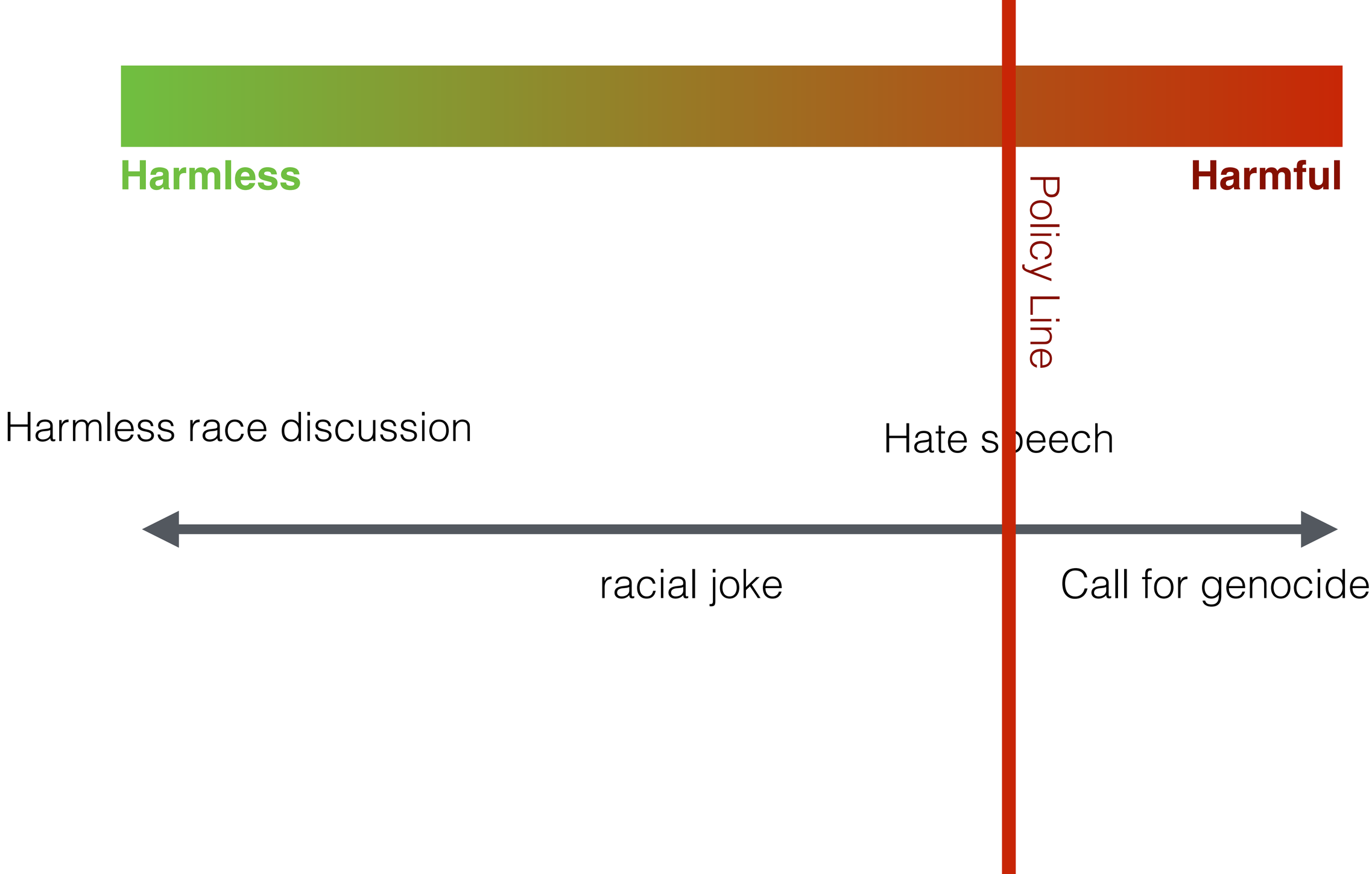
Harmless race discussion

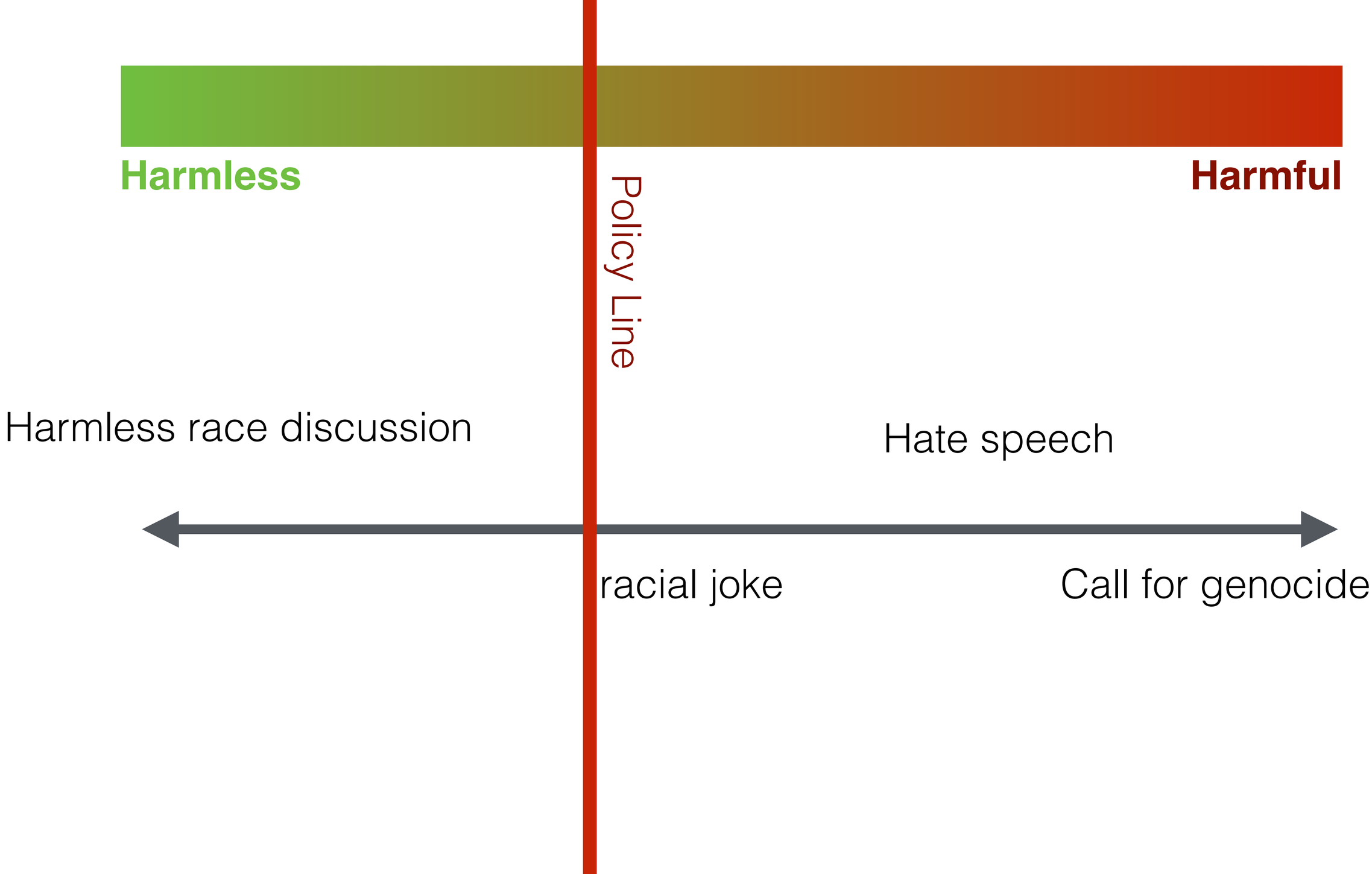
Hate speech

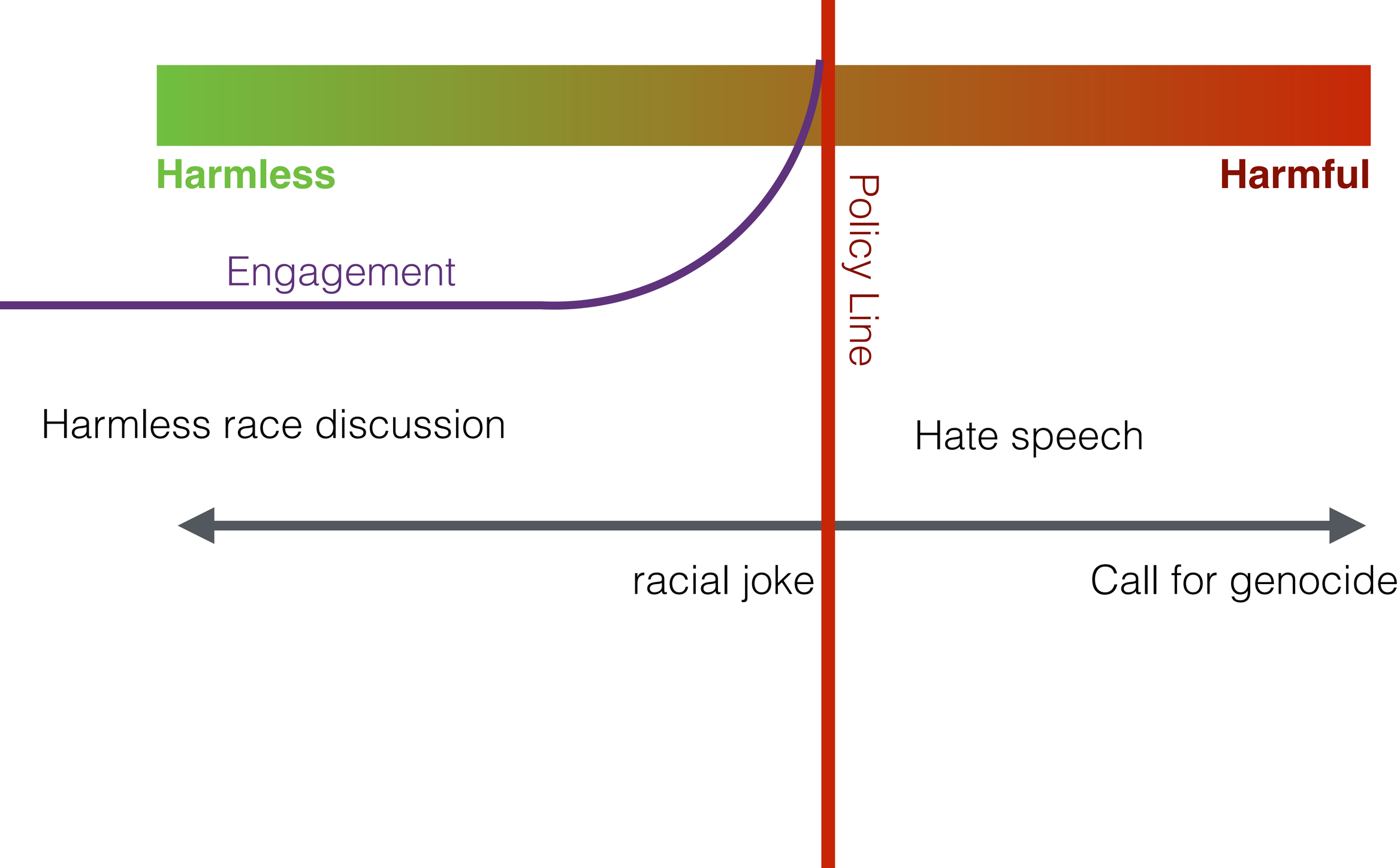
racial joke

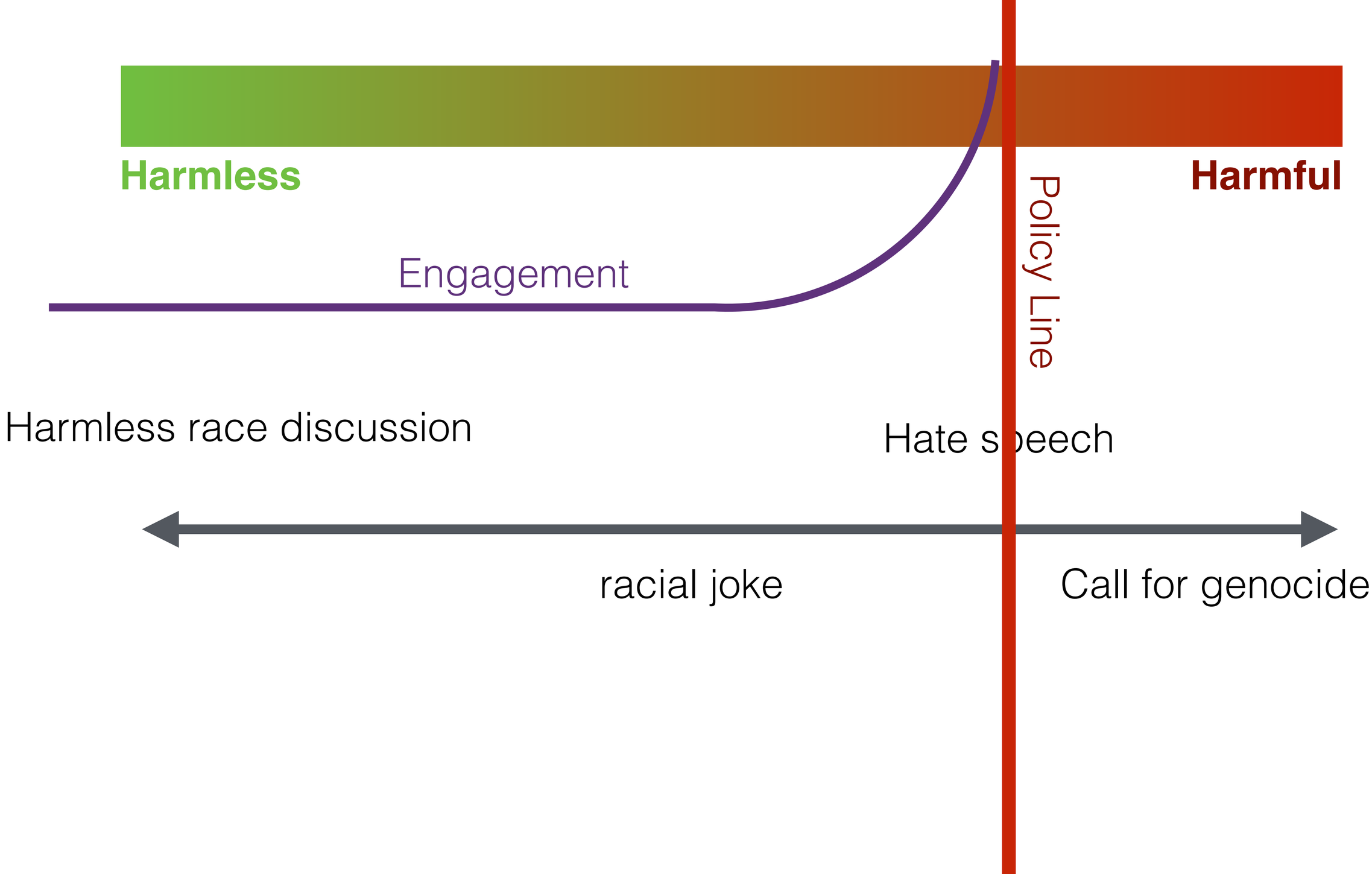
Call for genocide

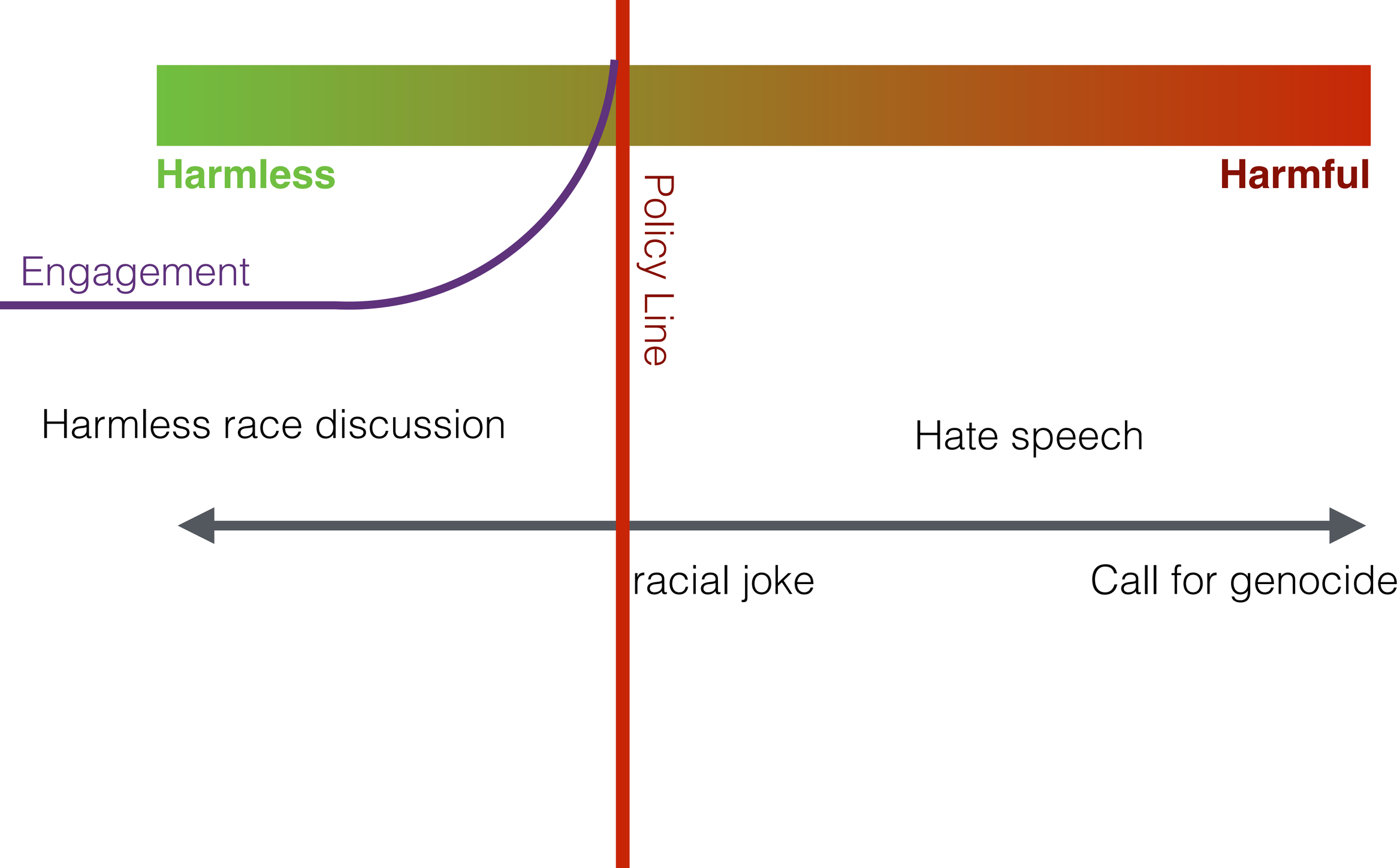

Harmless
Harmful
Policy Line
Harmless race discussion
Hate speech
racial joke
Call for genocide


Harmless
Harmful
Policy Line
Harmless race discussion
Hate speech
racial joke
Call for genocide

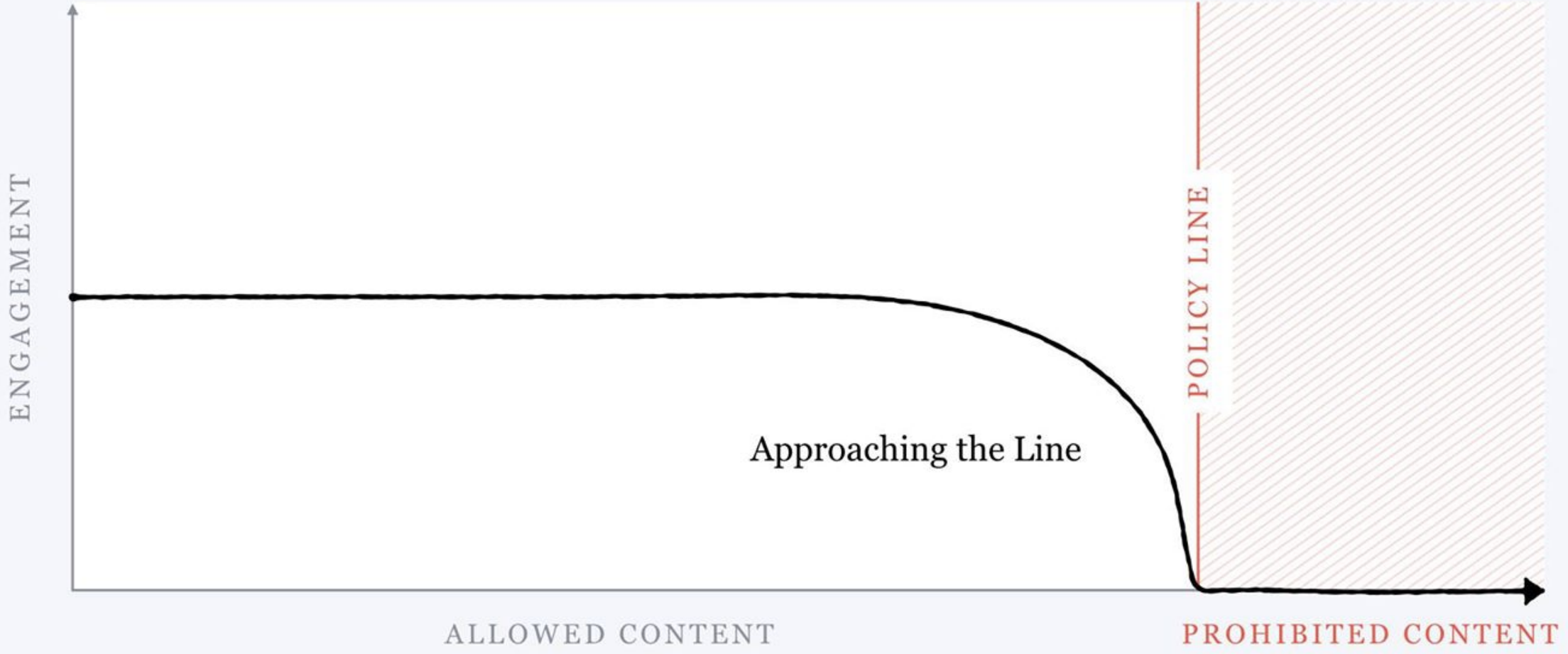

Harmless
Harmful
Policy Line
Harmless race discussion
Hate speech
racial joke
Call for genocide

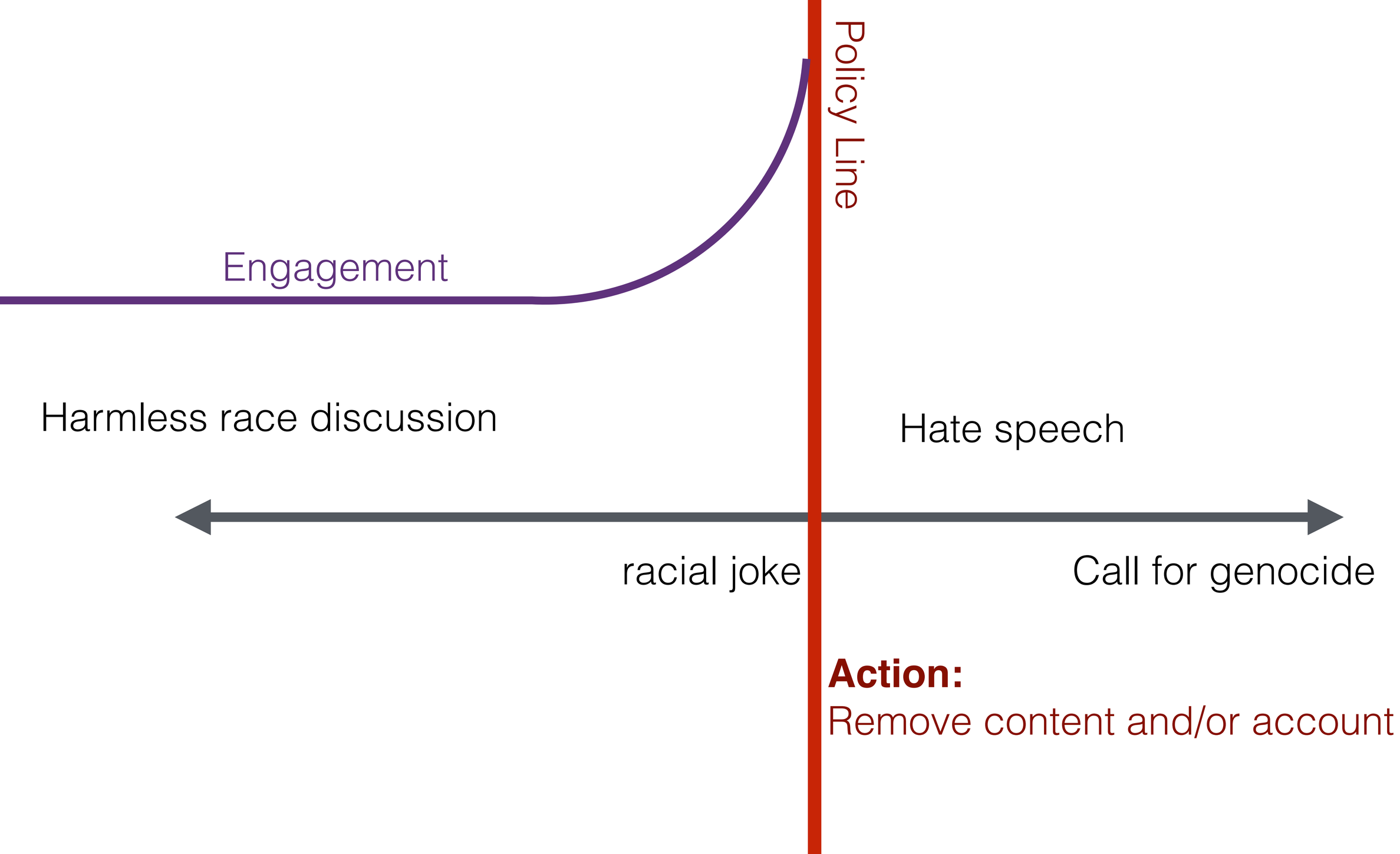

Harmless
Harmful
Policy Line
Engagement
Harmless race discussion
Hate speech
racial joke
Call for genocide

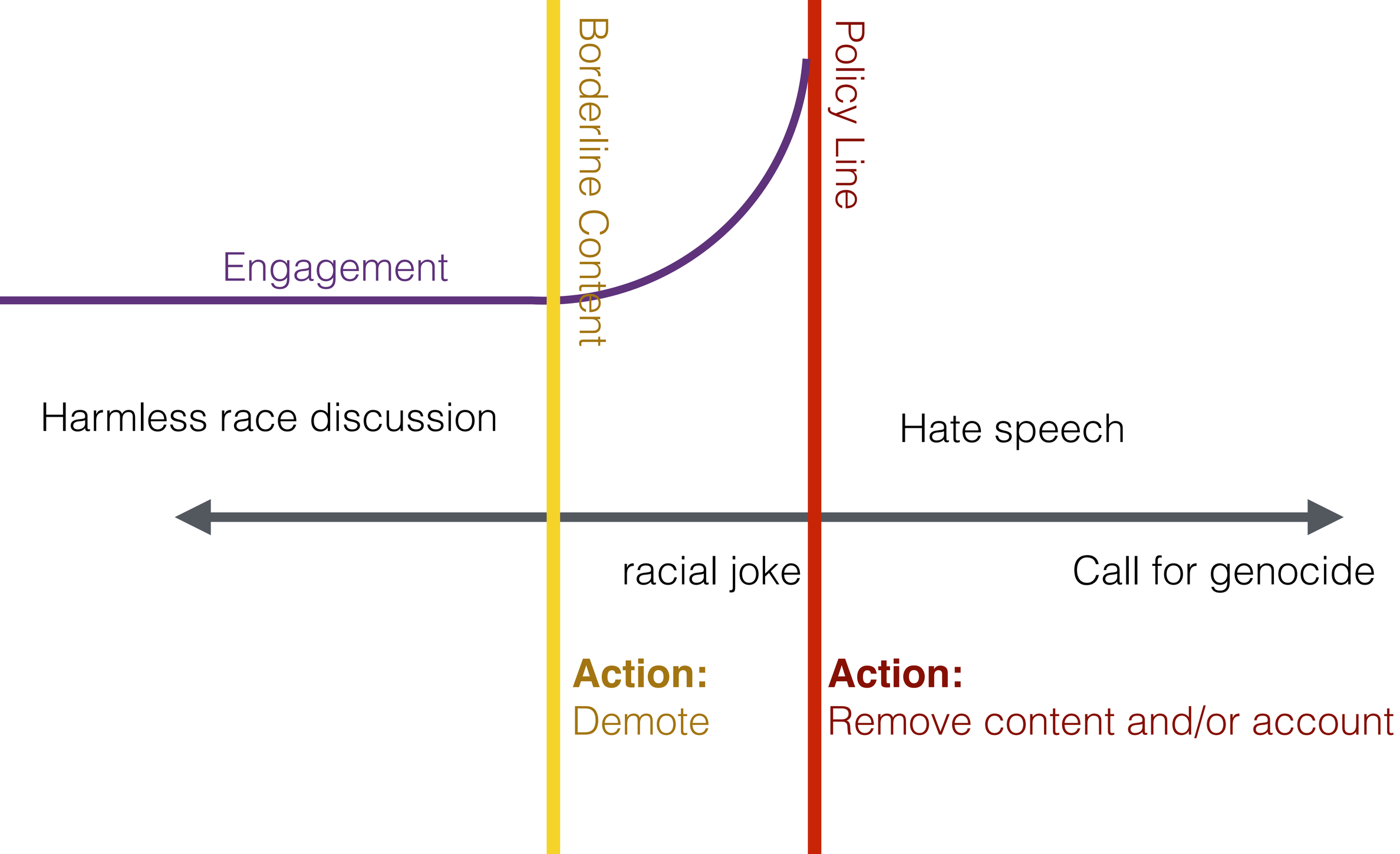

Harmless
Harmful
Policy Line
Engagement
Harmless race discussion
Hate speech
racial joke
Call for genocide

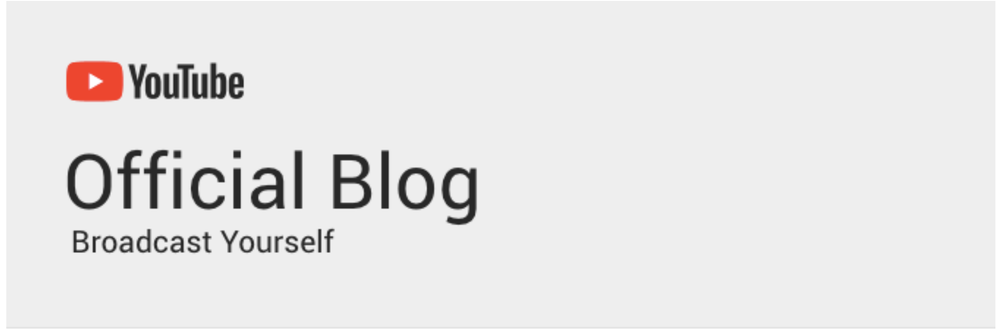

Harmless
Harmful
Engagement
Policy Line
Harmless race discussion
Hate speech
racial joke
Call for genocide

Adjusted to Discourage Borderline Content




Policy Line
Engagement
Harmless race discussion
Hate speech
racial joke
Call for genocide
Action:
Remove content and/or account


Borderline Content
Policy Line
Engagement
Harmless race discussion
Hate speech
racial joke
Call for genocide
Action:
Demote
Action:
Remove content and/or account

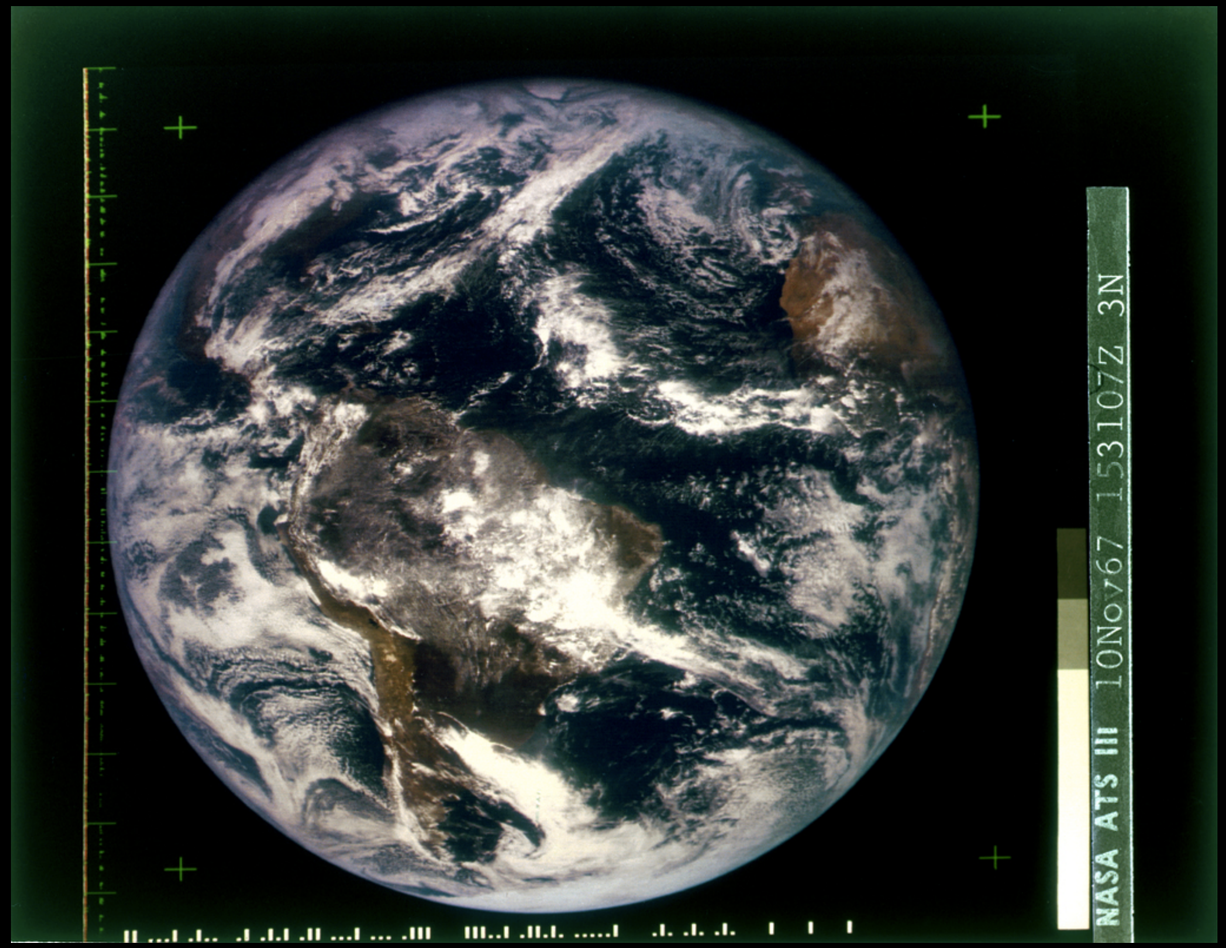


“we’ll begin reducing recommendations of borderline content and content that could misinform users in harmful ways.”

“such as:

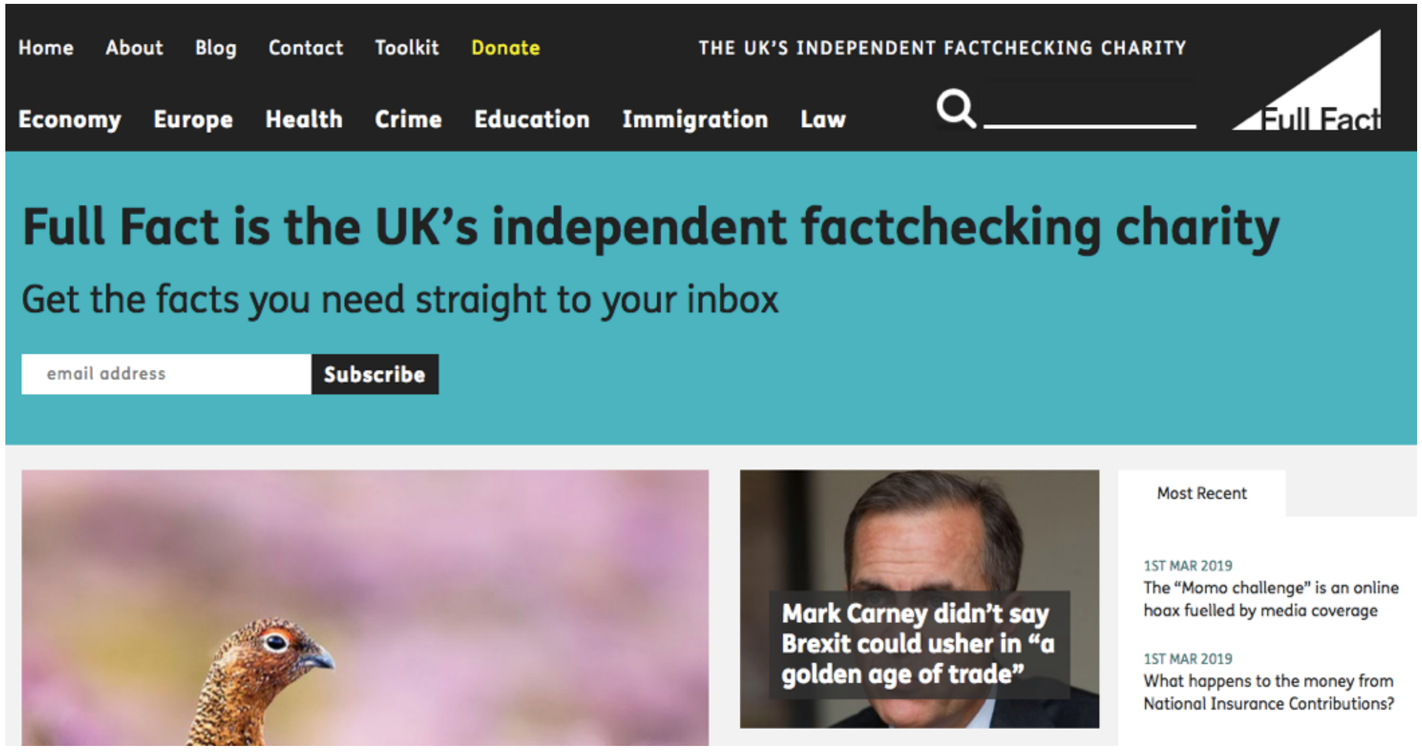
- videos promoting a phony miracle cure for a serious illness,
- claiming the earth is flat, or
- making blatantly false claims about historic events like 9/11”

https://youtube.googleblog.com/2019/01/continuing-our-work-to-improve.html

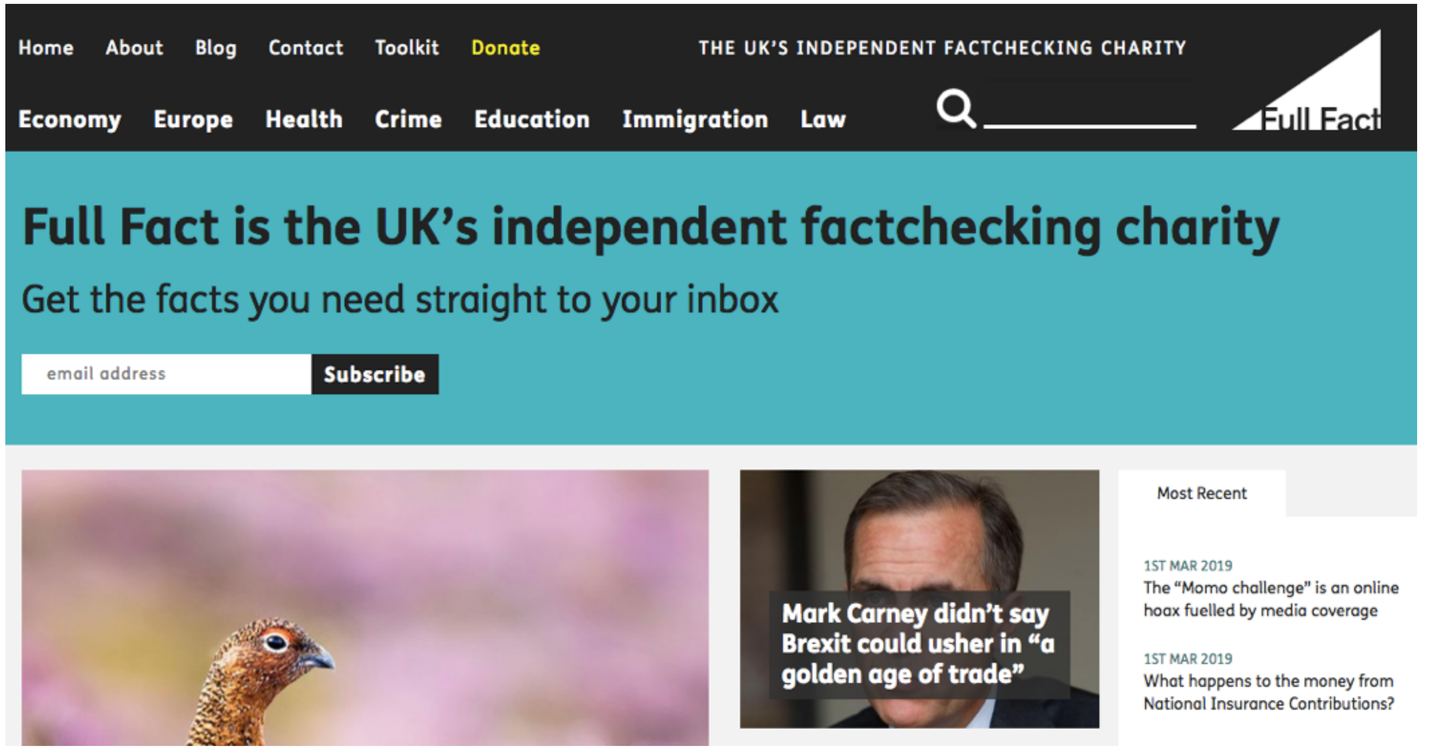


“Software is eating the world”
~Marc Andreessen

“You can't draw neat lines around planet-wide problems.”
~Dune

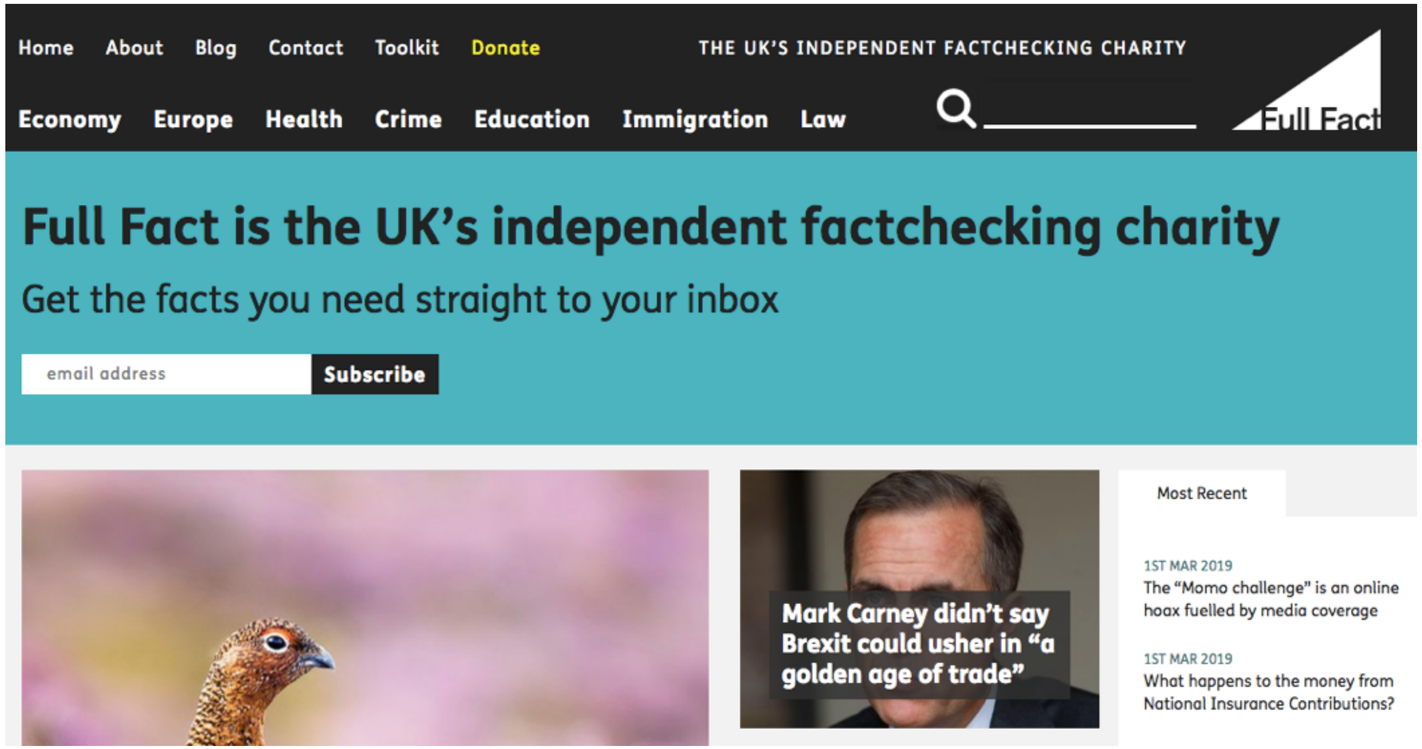


https://fullfact.org/



https://fullfact.org/

https://www.bbc.co.uk/news/technology-46836897

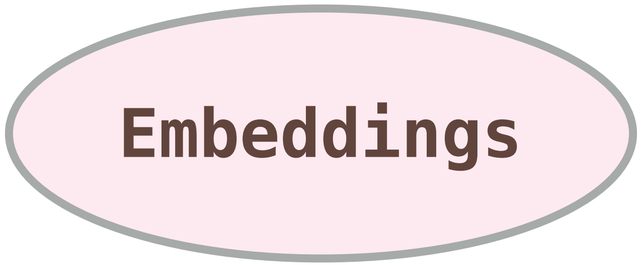


https://fullfact.org/

https://www.bbc.co.uk/news/technology-46836897

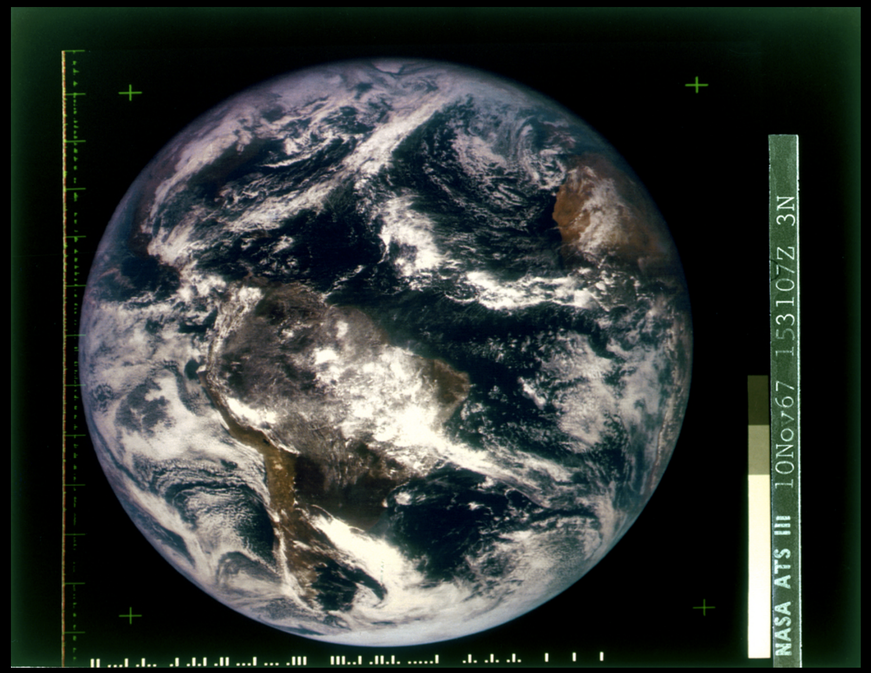
Sentence embeddings for automated factchecking - Lev Konstantinovskiy

https://www.youtube.com/watch?v=ddf0lgPCoSo



Sentence embeddings for automated factchecking - Lev Konstantinovskiy

https://www.youtube.com/watch?v=ddf0lgPCoSo



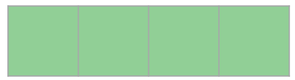
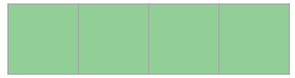
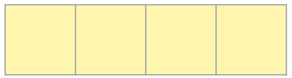
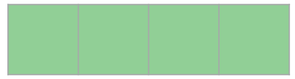
“Above all else, the mentat must be a generalist.

You can't draw neat lines around planet-wide problems.

The generalist looks outward; he looks for living principles, knowing full well that such principles change, that they develop. It is to the characteristics of change itself that the mentat-generalist must look. **There can be no permanent catalogue of such change, no handbook or manual.** You must look at it with as few preconceptions as possible, **asking yourself: "Now what is this thing doing?”**

~Dune

BACKUP

input word	output word	target	input • output	Softmax	Error
not 	thou	**1**	40	**0.95**	**0.05**
not 	aaron 	**0**	4	**0.04**	**-0.04**
not 	taco	**0**	2	**0.01**	**-0.01**

Deeper dive

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

1) Look up embeddings

2) Calculate Prediction

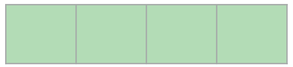
3) Project to Output Vocabulary

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

1) Look up embeddings

2) Calculate Prediction

3) Project to Output Vocabulary



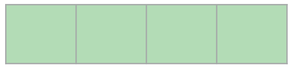
not

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

1) Look up embeddings

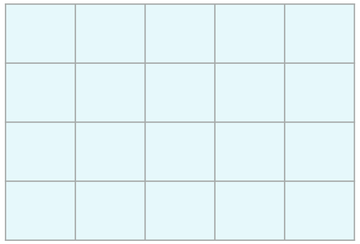
2) Calculate Prediction

3) Project to Output Vocabulary



not

X



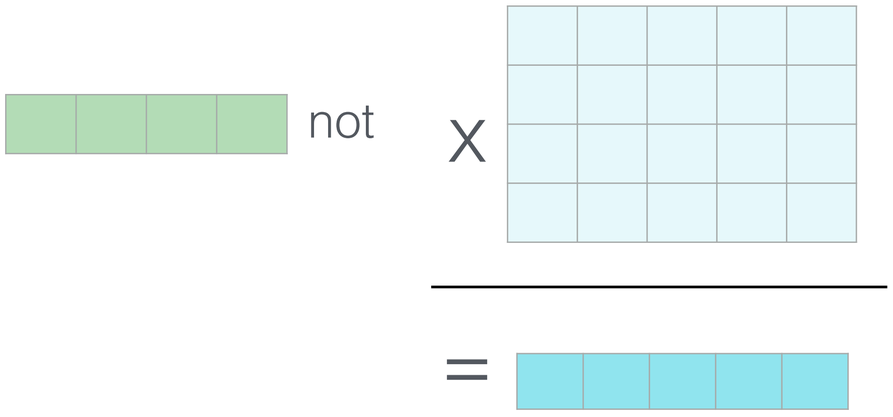
Hidden layer weights

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

1) Look up embeddings

2) Calculate Prediction

3) Project to Output Vocabulary



input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

1) Look up embeddings

not

2) Calculate Prediction

X

3) Project to Output Vocabulary

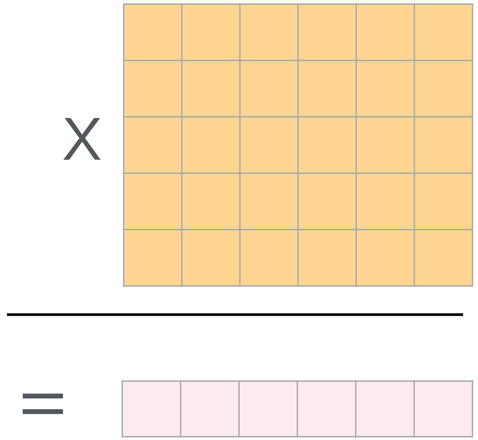
input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

1) Look up embeddings

not

2) Calculate Prediction

3) Project to Output Vocabulary

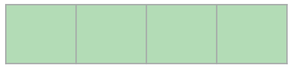


input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

1) Look up embeddings

2) Calculate Prediction

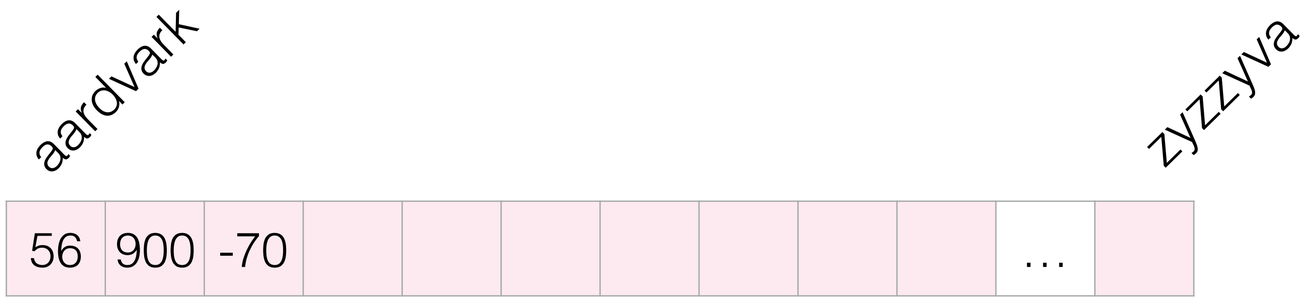
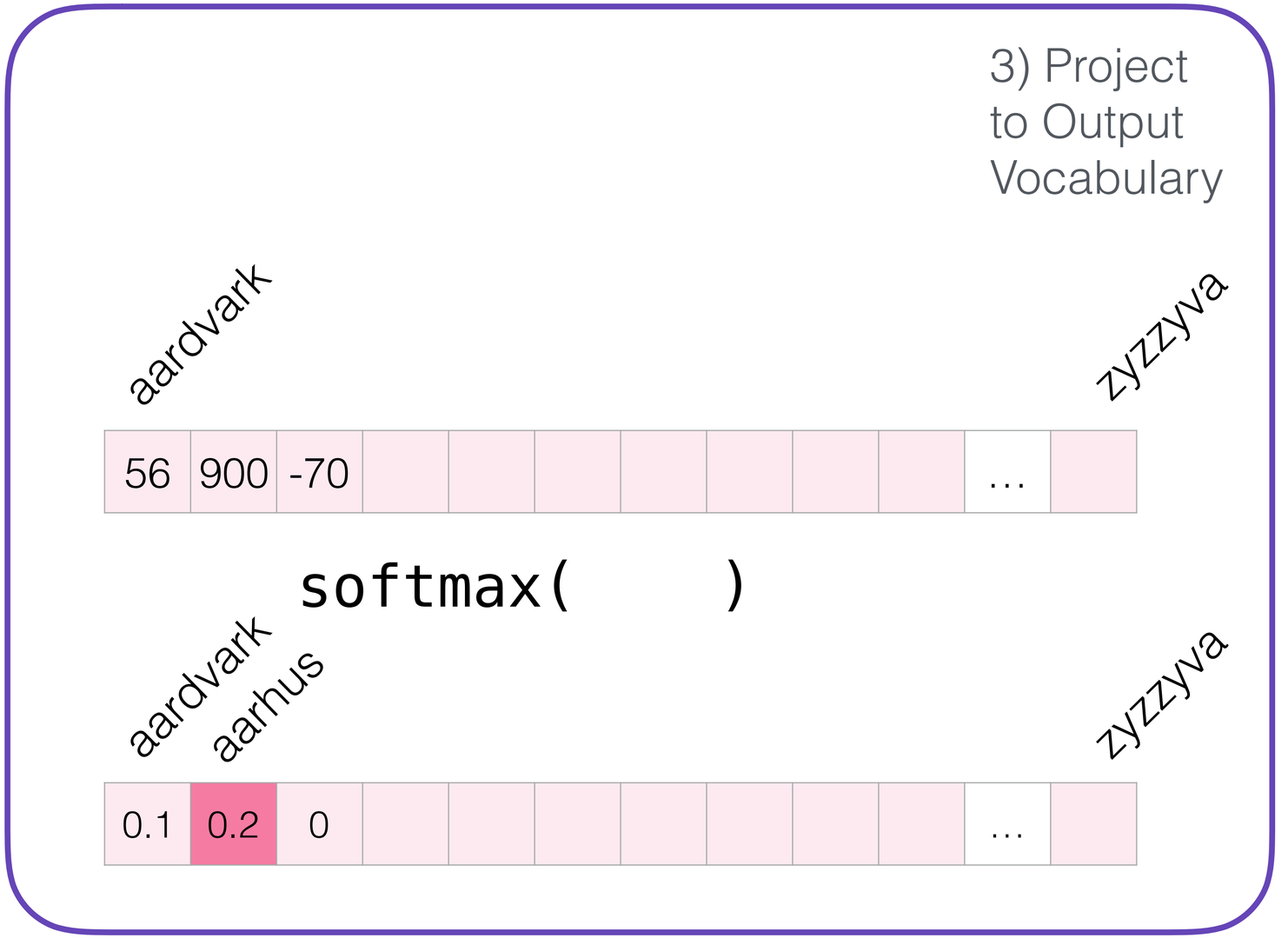
3) Project to Output Vocabulary



not

input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness

3) Project to Output Vocabulary



input word	target word
not	thou
not	shalt
not	make
not	a
make	shalt
make	not
make	a
make	machine
a	not
a	make
a	machine
a	in
machine	make
machine	a
machine	in
machine	the
in	a
in	machine
in	the
in	likeness