



How I convinced GPT to teach me about Hungarian folktales (without speaking a word of Hungarian)

This is both interesting and related to journalism, *I promise.*

Jonathan Soma

Knight Chair in Data Journalism, Columbia University

js4571@columbia.edu

@dangerscarf

<https://github.com/jsoma/mediaparty-folktales>

(Chat)GPT can answer questions

JS

What is the average weight of an adult capybara?



The average weight of an adult capybara, the largest rodent species in the world, is typically between 35 and 66 kilograms (77 and 146 pounds). However, individual capybaras can vary in size, and some exceptionally large individuals have been known to weigh up to 91 kilograms (201 pounds). It's worth noting that males are generally larger and heavier than females in this species.



EREDETI

NÉP MESÉK.

— 1862 —

ÖSSZEGYÜJTÖTTE

ARANY LÁSZLÓ.

— 1862 —

PEST.

KIADJA HECKENAST GUSZTÁV.

—

MDCCLXII.

707

fáradtak a lovak, megint elhajított a királyfi egy almát; harmadszor megint csak úgy tettek. Mikor már negyedszer is alig tudtak a lovak úszni, elértek szerencsésen a tulsó partot. Itt aztán gondolkozóba esett a királyfi, hogy most már hát itt volnának, de merre keressék az aranytollú madarat? Hiába gondolja megint csak el kellett szakítani ezt. Ott termett megint a róka.

— No róka pajtás — kérdezi — hol találjuk most már meg az aranytollú madarat?

— Hát látja-e felséged amott lámlani azt az arany palotát?

— Hogy ne látnám!

— No hát abban lakik az a királyfi, aki az aranytollú madár. Öltözzél ki a kaliczkád ingbe-gatyába csikósosan, azután hagyja itt, maga menjen oda a palota kertjébe, ott van a kaliczka az arany madárral egy arany fára felakasztva, azt vegye le, de vigyázzon, hogy valamelyik fához hozzá ne érjen, mert akkor az egész kert megzendül, a kertészek megfogják felségedet. Ha pedig mégis szerencsétlenül járna, megfognák, vitesse magát a királyhoz; majd a király kérdezi, hogy mert aholz a



Hungarian folktales

1862

madárhoz nyulni, csak ezt mondja felséged: „Hogy ne mertem volna, mikor széles ez világon, kis Magyarországon nincsen olyan betyár mint én.“ Többet aztán nem mondok.

Ugy tett a királyfi a hogy a róka mondta. Az imádat ott hagyta a tengerparton, maga felváltva a betyáros ruhába, elindult a palota felé. Nincs így ment egyenesen a palotába, hanem elkezdett alá; hát a mint oda ért, elállt szemei előtt még a hátulsó sövénye is a kertnek a supa aranyból-ezüstből volt fonva. A bámulta magát, bemászott szép csendes eritésen, szélylyel nézett a kertben, visszahúzta a szemét, hiszen nem látná meg az aranytollú madárt, minte meg nem látta volna. Elment a palota előtt volna is, úgy elvette szemét, mint a gyerekek a nagy csillagás-villogás; de hogy ismeretlen volt a királyfi! mikor meglátta, fának ezüst volt a törzsé, arany a galamánya, a gyümölcsé. Mikor aztán félíg-mélyig a kertben elindult a madarat keresni, meg is találta nemsokára, ott ugrált egy gyémánt kaliczkában, egy ágas-bogas arany fára felakasztva. A mint meglátta a királyfi, ment egyenesen érte, hogy levegye; nem volt magasra téve, felérte a földről is, nem kellett érte felmászni, hanem a mint vette vol-

GPT has *not* read this book.

How can we teach it?

"Fine-tune" or "train" a new model?

ion API reference Examples

Fine-tuning

Learn how to customize a model for your application.

Introduction

Fine-tuning lets you get more out of the models available through the API by providing:

- 1 Higher quality results than prompt design
- 2 Ability to train on more examples than can fit in a prompt
- 3 Token savings due to shorter prompts
- 4 Lower latency requests

GPT-3 has been pre-trained on a vast amount of text from the open internet. When given a prompt with just a few examples, it can often intuit what task you are trying to perform and generate a plausible completion. This is often called "few-shot learning."

Fine-tuning improves on few-shot learning by training on many more examples than can fit in the prompt, letting you achieve better results on a wide number of tasks. **Once a model has been fine-tuned, you won't need to provide examples in the prompt anymore.** This saves costs and enables lower-latency requests.

3.17564v2 - May 2023

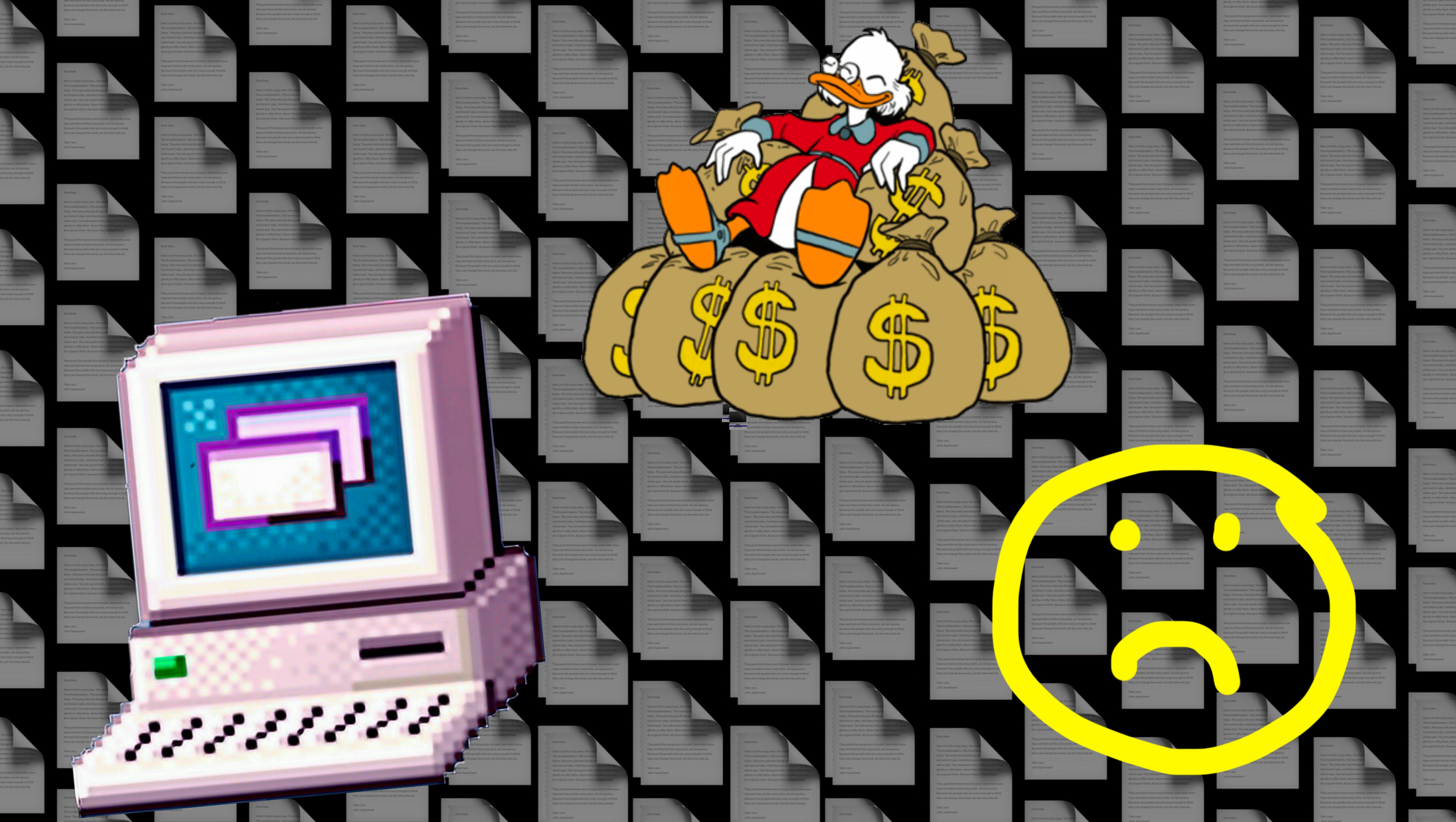
Contents

BloombergGPT: A Large Language Model for Finance

Shijie Wu^{1,*}, Ozan Irsoy^{1,*}, Steven Lu^{1,*}, Vadim Dabrowski¹, Mark Dredze¹, Sebastian Gehrmann¹, Prabhanjan Kambadur¹, David Rosenberg², Gideon M. Bloomberg, New York, NY USA
² Bloomberg, Toronto, ON Canada
³ Computer Science, Johns Hopkins University, Baltimore, MD USA

Abstract

The use of NLP in the realm of financial technology is broad and complex, with applications ranging from sentiment analysis and named entity recognition to question answering. Large Language Models (LLMs) have been shown to be effective on a variety of tasks; however, no LLM specialized for the financial domain has been reported in literature. In this work, we present BLOOMBERGGPT, a 50 billion parameter language model that is trained on a wide range of financial data. We construct a 363 billion token dataset based on Bloomberg's 345 billion tokens from general purpose datasets. We validate BLOOMBERGGPT on standard LLM benchmarks, open financial benchmarks, and a suite of internal benchmarks that most accurately reflect our intended usage. Our mixed dataset training leads to a model that outperforms existing models on financial tasks by significant margins without sacrificing performance on general LLM benchmarks. Additionally, we explain our modeling choices, training process, and evaluation methodology. We release Training Chronicle (Appendix C) detailing our experience in training BLOOMBERGGPT.



In-context learning

JS

The devil owns sixteen golden cabbages, and gloves of the skin of a fish.

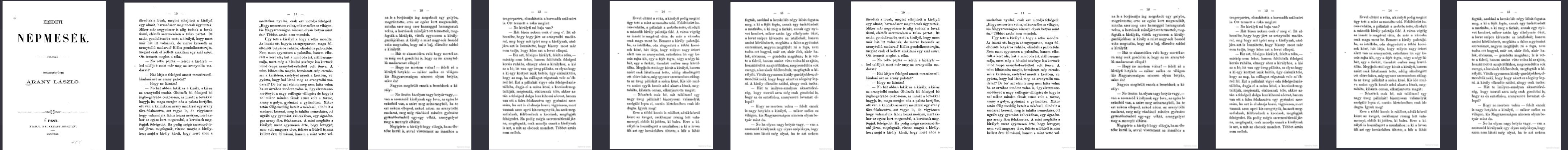
How many golden cabbages does the devil own? Answer using the information provided above.



The devil owns 16 golden cabbages.

Let's just send the entire book!!!!

JS



How many golden cabbages does the devil own? Answer using the information provided above.



The devil owns 16 golden cabbages.

Let's just send relevant passages

- 11 -

madárhoz nyulni, csak ezt mondja felséged:
„Hogy ne mertem volna, mikor szélesez világom,
kis Magyarországon nincsen olyan betyár mint
én.“ Többet aztán nem mondok.
Ugy tett a királyfi a hogy a róka mondta.

Az inasát ott hagyta a tengerparton, maga fel-
öltözött betyáros ruhába, elindult a palota felé.
Nem ment egyenesen a palotába, hanem elke-
rült a kert alá; hát a mint oda ért, elállt szeme-
szája, mert még a hárulsó sövénye is a kertnek
mind csupa aranyból-ezüstből volt fonva. A
mint kibámulta magát, bemászott szép csende-
sen a kerítésen, szélylyel nézett a kertben, vi-
gyázta, hogy hol látná meg az aranytollú ma-
darat? De biz' azt eleinte meg nem láttá volna
ha az orrához ütődött volna is, úgy elvette sze-
m-fényét a nagy csillagás-villogás; de hogy is
ne! mikor minden fának ezüst volt a törzsé,
arany a galva, gyémánt a gyümölcsé. Mikor
aztán félíg-meddig betelt a nézéssel, elindult a
madarakat keresni, meg is találta nemsokára, ott
ugrált egy gyémánt kaliczkában, egy ágas-bo-
gas arany fára felakaszta. A mint megláttá a
királyfi, ment egyenesen érte, hogy levegye;
nem volt magasra téve, felérte a földről is, nem
kellett érte felmászni, hanem a mint vette vol-

- 18 -

Kapott a királyfi két kézzel a tanácon,
mindjárt elhatározta, hogy még azon az éj-
szakán, majd ha mindenki lenyugszik a kirá-
lyi udvarban, megsöknek, reggelre pedig, ha
lehet, — túl teszik magukat a határon. Evvel
a királyfi kiment a kissasszony szobájából, vár-
tak a szomszéd királynak, hogy ne-
cél; csak akkor jutott eszébe, hogy ne-
megtartotta a szomszéd királynak kell

- 19 -

akármint lesz, el kell neki lopni a kisasszonyt,
jobb hát ha most elszökteti, azután majd taná-
csot kérhet a rókától.

El is szöktette szerencsésen, kimentek a me-
zőre a királyfi inasához, ott, — mig az inas a
lovakat nyergelte, — a királyfi félre ment, el-
szakította az ötödik szál szört is. Ott termett
a róka:

— Mi a baj felséges királyfi?

— Hát bizon most az, hogy nem tudom a
királykisasszontól váljak-e meg, vagy madár
nélkül menjek haza? — mert így, meg így áll
a dolog.

— Sohse' búsuljon azon a királyfi, majd
elintézem én azt, most majd együtt megyek fel-
ségeddel.

Mentek aztán együtt négyen: a királyfi, a
szép királykisasszony, az inas meg a róka. Mi-
kor ahoz a városhoz értek, ahova a kisasszonyt
kellett volna vinni, a város végén a róka épe-
olyan szép kisasszonyná változott mint az igá-
zi, úgy hogy a királyfi is alig tudta megismerni
őket, a róka-kisasszony mondta az igazinak, hog-
csak menjenek az inassal, nekik most a király
val egy kis dolguk van, hanem majd után
mennek. Ment aztán a kisasszony tovább, a

Search for matching words...

What did Zsuzska steal from the devil?

Zsuzska just laughed, but when the laughter didn't turn into tears, because the **devil** was chasing after her.

So they went to the king again, they lied to him that **Zsuzska** had said that the **devil** had a bag of golden nuts, and he could **steal** that too.

...but that doesn't always work.

What did Zsuzska steal from the devil?

Zsuzska, you dog, you killed my three beautiful daughters, **stole** my sea-walking shoes, **stole** my sea-beating staff, now you're **taking** my golden-headed cabbage!



cat



tiger



dog



cat



tiger



dog



wolf

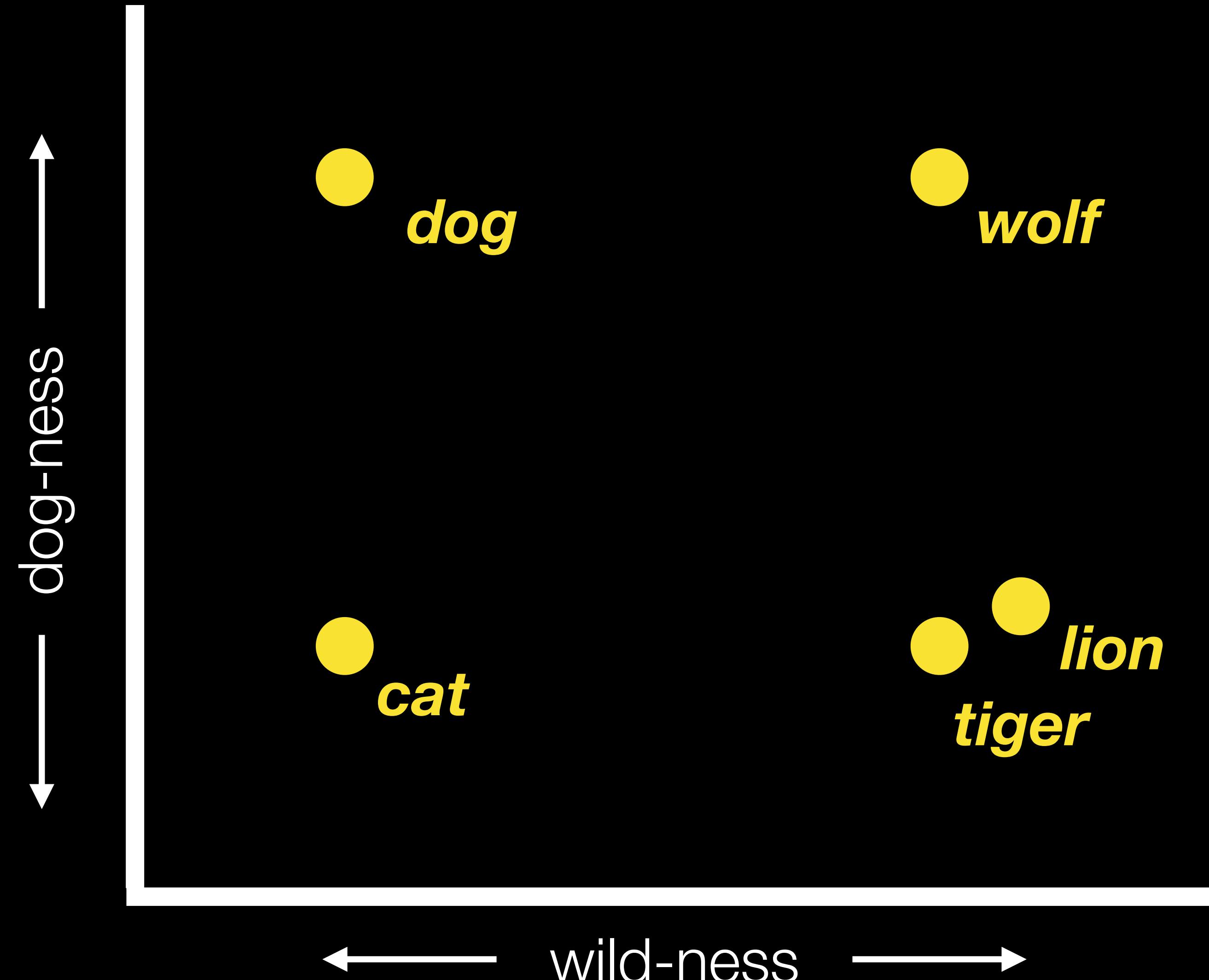


cat

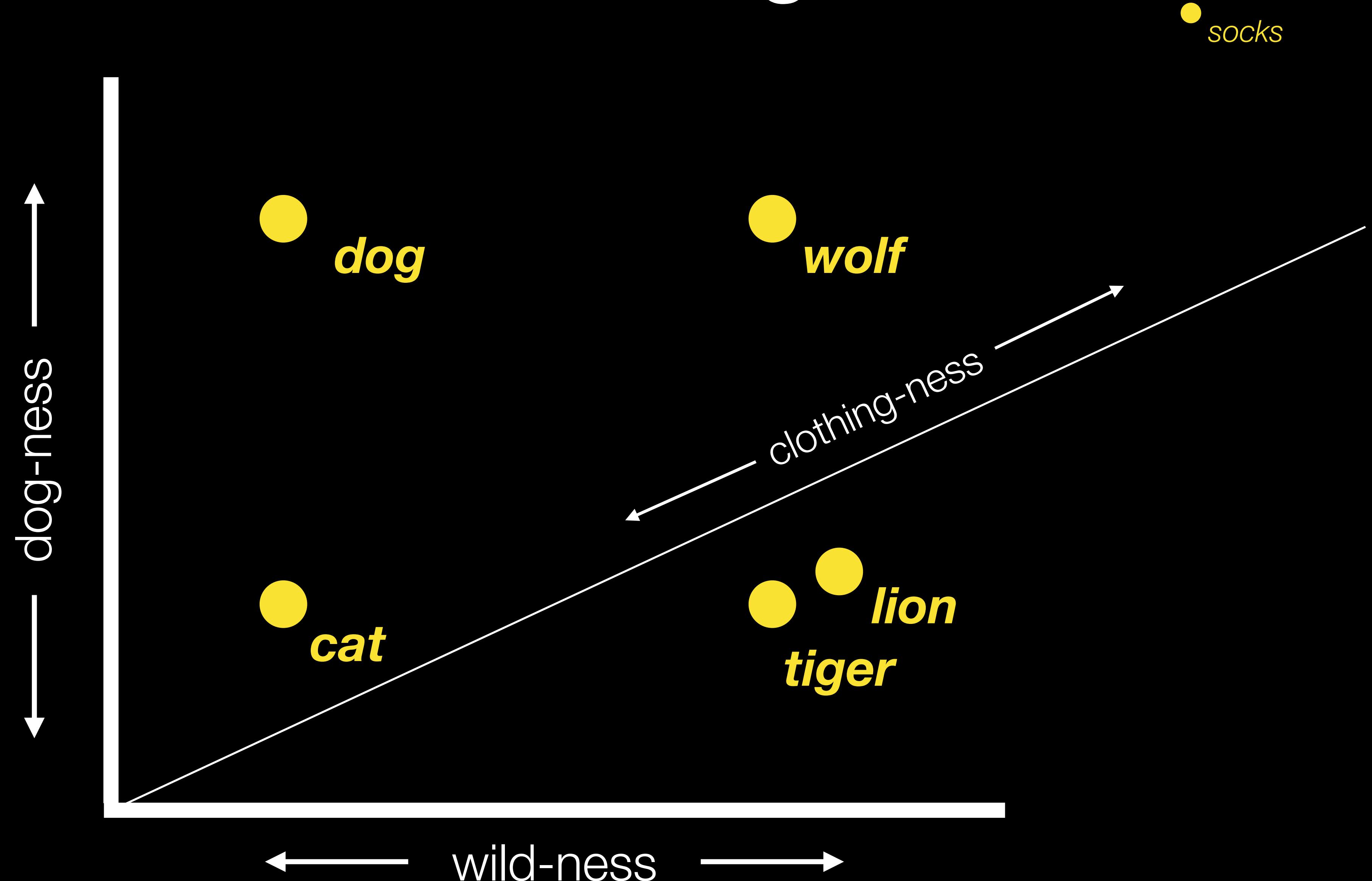


tiger

Word embeddings



Word embeddings



Word embeddings

	dog	wild	clothing	fuzzy	code
<i>cat</i>	1	3	1	4	5
<i>dog</i>	10	2	0	6	0
<i>wolf</i>	9	9	1	2	0
<i>lion</i>	2	0	1	6	2
<i>tiger</i>	1	0	2	3	3
<i>socks</i>	0	2	9	7	1

Searching for similar words



Searching for similar words paragraphs

• • •
kings being grumpy

• •
cats singing songs

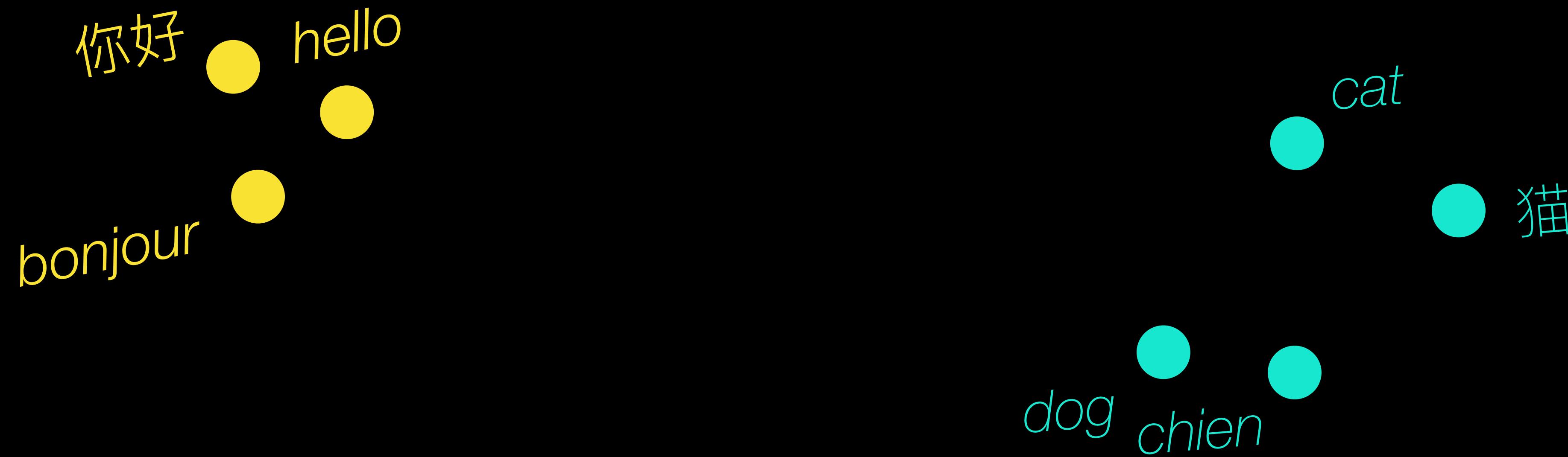
• • •
things being stolen

from kings

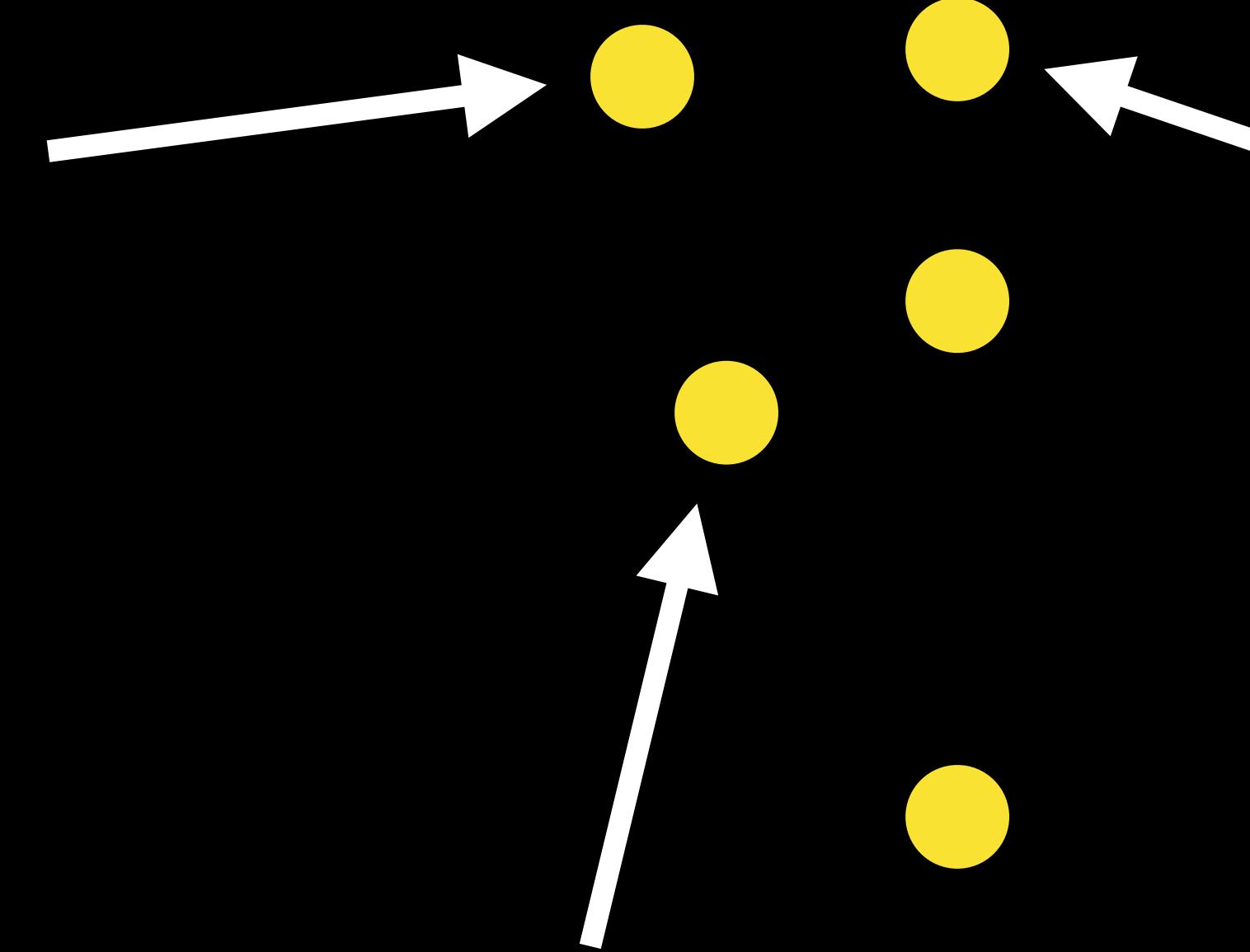
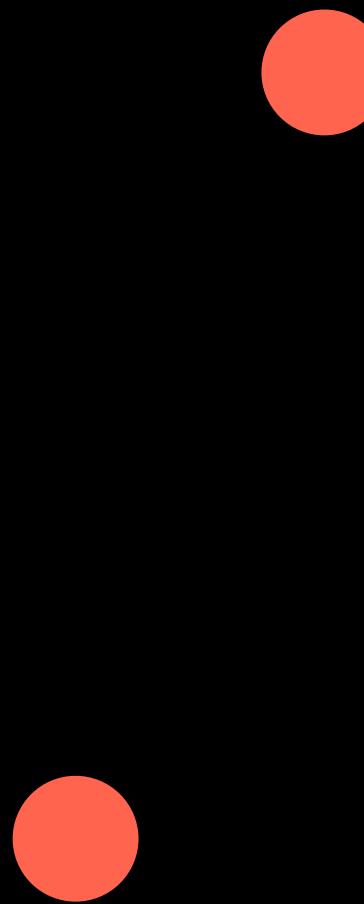
• • •
from the devil

„Hogy”
kis Magyarorszagon
én.” Többet aztán nem mondott.
Ugy tett a királyfi a hogy a róka monda felé.
Az inasát ott hagyta a tengerparton, maga fel-
öltözött betyáros ruhába, elindult a palota felé.
Nem ment egyenesen a palotába, hanem elke-
rült a kert alá; hát a mint oda ért, elállt szeme-
szája, mert még a hátulsó sövénye is a kertnek
mind csupa aranyból-eziüstből volt fonva. A
mint kibámulta magát, bemászott szép csende-
sen a kerítésen, szélylyel nézett a kertben, vi-
gyázta, hogy hol látná meg az aranytollú ma-
darat? De biz' azt eleinte meg nem látta volna
ha az orrához ütődött volna is, úgy elvette sze-
me-fényét a nagy csillogás-villogás; de hogy is
ne! mikor minden fának eziüst volt a törzse,
arany a galya, gyémánt a gyümölcs. Mikor
ig-meddig betelt a nézéssel, elindult a
is találta nemsokára, ott
egy ágas-bo-

Multi-language embeddings



"What did Zsuzska steal from the devil?"



"Hiába tagadta szegény Zsuzska, nem használt semmit, elindult hát nagyszomorúan. Épen éjfél volt, mikor az ördög házához ért, aludt az ördögis, a felesége is. Zsuzska csendesen belopódzott, ellopta a tenger-ütőpálczát, avval bekiáltott az ablakon. – Hej ördög, viszem ám szép lyányomat, elloptad a tenger-lépő cipőmet, most viszed a tenger-ütő pálczámat, de majd meglakolsz te ezért. Utána is szaladt, de megint csak a tengerparton tudott közel jutni hozzá, ott meg Zsuzska megütötte a tengert a tenger-ütő pálczával, kétfelé vált előtte, utána ment egyenesen a királyhoz."

De Zsuzska nem adta: „Tán bolond vagyok, hogy visszaadjam, mikor kívülvagyok már vele az udvaron?!” Az ördög kergette egy darabig, de seholgye tudta utolérni, utoljára is visszafordult, Zsuzska pedig mentegyenesen a király elibe, od'adta neki az arany fej káposztát. – No felséges király elhoztam már ezt is! A két nénjét, Zsuzskának, majd hogy meg nem ütötte a guta, másnap megintbementek a királyhoz. – Jaj felséges király van még annak az ördögnek egy arany kis gyermekes király fölött. Megint behivatta a beszéli fűnek-fának, hogy ő azt is eltudná lopni. Megint behivatta a király Zsuzskát. – Fiam Zsuzska, azt hallottam, hogy van annak az ördögnek egy arany kisgyermek is, arany bőlcsoben, te azt is el tudod lopni, azt beszélted, azért ha az éjjel el nem lopod, halálnak halálával halsz meg.

read

split

embed

filter

query

Nem ment egyenesen a palotába, hanem elkerült a kert alá; hát a mint oda ért, elállt szemeszaja, mert még a hátulsó sóvényle is a kertnek

PEST.
KIADJA HECKENAST GUSZTÁV.
MDCCCLXII.

Digitized by Google

— 11 —

madárhoz nyulni, csak ezt mondja felséged: „Hogy ne mertem volna, mikor széles ez világom, kis Magyarországon nincsen olyan betyár mint én.“ Többet aztán nem mondok.

Ugy tett a királyfi a hogy a róka mondta. Az inasát ott hagyta a tengerparton, maga felőltözített betyáros ruhába, elindult palota felé. Nem ment egyenesen a palotába, hanem elkerült a kert alá; hát a mint oda ért, elállt szemeszaja, mert még a hátulsó sóvényle is a kertnek mind csupa aranyból-ezüstből volt fonva. A mint kibámulta magát, bemászott szép csenderesen a kerítésen, szélyivel nézett a kertben, vigázta, hogy hol látná meg az aranytollú madarat? De biz' azt eleinte meg nem láttá volna ha az orrához ütődött volna is, úgy elvette szeme-fényét a nagy csillagás-villogás; de hogy is

mind csupa aranyból-ezüstből volt fonva. A mint kibámulta magát, bemászott szép csenderesen a kerítésen, szélyivel nézett a kertben, vigázta, hogy hol látná meg az aranytollú madarat? De biz' azt eleinte meg nem láttá volna ha az orrához ütődött volna is, úgy elvette szeme-fényét a nagy csillagás-villogás; de hogy is

ne! mikor minden fának ezüst volt a törzse, arany a galya, gyémánt a gyümölcs. Mikor aztán félíg-meddig betelt a nézéssel, elindult a madarat keresni, meg is találta nemsokára, ott ugrált egy gyémánt kaliczkában, egy ágas-bo-

gas arany fára felakaszta. A mint meglátta a királyfi, ment egyenesen érte, hogy levegye; nem volt magasra téve, felérte a földről is, nem kellett érte felmászni, hanem a mint vette vol-

— 10 —

fáradtak a lovak, megint elhajított a királyfi egy almát; harmadszor megint csak úgy tettek. Mikor már negyedszer is alig tudtak a lovak úszni, elértek szerencsén a tulsó partot. Itt aztán gondolkozóba esett a királyfi, hogy most már hát itt volnának, de merre keressék az aranytollú madarat? Hiába gondolkozott rajta, megint csak el kellett szakítani egy szál szört. Ott termett megint a róka.

— 11 —

madárhoz nyulni, csak ezt mondja felséged: „Hogy ne mertem volna, mikor széles ez világom, kis Magyarországon nincsen olyan betyár mint én.“ Többet aztán nem mondok.

Ugy tett a királyfi a hogy a róka mondta. Az inasát ott hagyta a tengerparton, maga felőltözített betyáros ruhába, elindult palota felé. Nem ment egyenesen a palotába, hanem elkerült a kert alá; hát a mint oda ért, elállt szemeszaja, mert még a hátulsó sóvényle is a kertnek mind csupa aranyból-ezüstből volt fonva. A mint kibámulta magát, bemászott szép csenderesen a kerítésen, szélyivel nézett a kertben, vigázta, hogy hol látná meg az aranytollú madarat? De biz' azt eleinte meg nem láttá volna ha az orrához ütődött volna is, úgy elvette szeme-fényét a nagy csillagás-villogás; de hogy is

— No róka pajtás — kérde a királyfi — hol találjuk most már meg az aranytollu madarat?

— Hát látja-e felséged amott messzire csillámlani azt az arany palotát?

— Hogy ne látnám!

— No hát abban lakik az a király, a kié az az aranytollú madár. Öltözék fel felséged bő ingbe-gatyába csikósan, az inasát a lovakkal hagyja itt, maga menjen oda a palota kertjébe, ott van a kaliczka az arany madárral egy arany fára felakaszta, azt vegye le, de vigyázzon hogy valamelyik fához hozzá ne érjen, mert akkor az egész kert megzendül, a kertések megfogják felségedet. Ha pedig mégis szerencsétlenül járna, megfognák, vitesse magát a királyhoz; majd a király kérde, hogy mert aholhoz a

1.6, 1.4, 8.3, ...

0.7, 9.0, 6.5, ...

4.3, 2.7, 8.9, ...

1.1, 6.7, 3.2, ...

7.8, 5.4, 0.3, ...

2.8, 9.3, 4.6, ...

1.8, 3.4, 6.1, ...

0.2, 8.6, 5.0, ...

5.2, 3.9, 2.1, ...

7.6, 1.4, 8.3, ...

Nem ment egyenesen a palotába, hanem elkerült a kert alá; hát a mint oda ért, elállt szemeszaja, mert még a hátulsó sóvényle is a kertnek

mind csupa aranyból-ezüstből volt fonva. A mint kibámulta magát, bemászott szép csenderesen a kerítésen, szélyivel nézett a kertben, vigázta, hogy hol látná meg az aranytollú madarat? De biz' azt eleinte meg nem láttá volna ha az orrához ütődött volna is, úgy elvette szeme-fényét a nagy csillagás-villogás; de hogy is

ne! mikor minden fának ezüst volt a törzse, arany a galya, gyémánt a gyümölcs. Mikor aztán félíg-meddig betelt a nézéssel, elindult a madarat keresni, meg is találta nemsokára, ott ugrált egy gyémánt kaliczkában, egy ágas-bo-

ne! mikor minden fának ezüst volt a törzse, arany a galya, gyémánt a gyümölcs. Mikor aztán félíg-meddig betelt a nézéssel, elindult a madarat keresni, meg is találta nemsokára, ott ugrált egy gyémánt kaliczkában, egy ágas-bo-

— 10 —

fáradtak a lovak, megint elhajított a királyfi egy almát; harmadszor megint csak úgy tettek. Mikor már negyedszer is alig tudtak a lovak úszni, elértek szerencsén a tulsó partot. Itt aztán gondolkozóba esett a királyfi, hogy most már hát itt volnának, de merre keressék az aranytollú madarat? Hiába gondolkozott rajta, megint csak el kellett szakítani egy szál szört. Ott termett megint a róka.

— No róka pajtás — kérde a királyfi — hol találjuk most már meg az aranytollu madarat?

— Hát látja-e felséged amott messzire csillámlani azt az arany palotát?

— Hogy ne látnám!

+ question



Using the texts below, answer the question "what did Zsuzska steal from the devil?"



Hiába tagadta szegény Zsuzska, nem használt semmit, elindult hát nagyszomorúan. Épen éjfél volt, mikor az ördög házához ért, aludt az ördögis, a felesége is. Zsuzska csendesen belopódzott, ellopta a tenger-ütőpálczát, avval bekiáltott az ablakon. – Hej ördög, viszem ám már a tenger-ütő pálczádat is. – Hej kutya Zsuzska, megölettek három szép lyányomat, elloptad atenger-lépő czipőmet, most viszed a tenger-ütő pálczámat, de majd meglakolsz te ezért. Utána is szaladt, de megint csak a tengerparton tudott közel jutni hozzá, ott meg Zsuzska megütötte a tengert a tenger-ütő pálczával, kétfelé vált előtte, utána meg összecsapódott, megint nem foghatta megaz ördög. Zsuzska ment egyenesen a királyhoz.

De Zsuzska nem adta; „Tán bolond vagyok, hogy visszaadjam, mikor kivülvagyok már vele az udvaron?!” Az ördög kergette egy darabig, de sehogyse tudta utolérni, utoljára is visszafordult, Zsuzska pedig mentegyenesen a király elibe, od’adta neki az arany fej káposztát. – No felséges király elhoztam már ezt is! A két nénjét Zsuzskának, majd hogy meg nem ütötte a guta, mikormegtudták, hogy Zsuzskának most se’ lett semmi baja, másnap megint bementek a királyhoz. – Jaj felséges király van még annak az ördögnek egy arany kis gyermekelis arany bölcsőben, Zsuzska azt beszéli fűnek-fának, hogy ő azt is eltudná lopni. Megint behívatta a király Zsuzskát. – Fiam Zsuzska, azt hallottam, hogy van annak az ördögnek egy arany kisgyermeke is, arany bölcsőben, te azt is el tudod lopni, azt beszélted, azért ha az éjjel el nem lopod, halálnak halálával halsz meg.

– No felséges király, elhoztam már a tengerütő pálczát is. A király még jobban megszerette Zsuzskát, hogy olyan életre való, de anénjei még jobban irigykedtek rá, csakhamar megint avval árulták be, hogy van annak az ördögnek egy arany fej káposztája is, Zsuzska azt is el tudná lopni, azt mondta. A király megint ráparancsolt Zsuzskára erősparancsolattal, hogy ha a káposztát el nem lopja, halálnak halálával hal meg. Elindult hát szegény Zsuzska megint, el is ért szerencsésen épen éjfélre az ördög kertjibe, levágta az arany fej káposztát, avval bekiáltott az ablakon. – Hej ördög, viszem ám már az arany fej káposztádat is. – Hej kutya Zsuzska, megölettek három szép lyányomat, elloptad atenger-lépő czipőmet, elloptad a tenger-ütő pálczámat, most viszed az arany fej káposztámat, csak ezt az egyet add vissza, soha szemedre sevetem.

Using "what" to answer the question "what?"



```
query = "What did Zsuzská steal from the devil?"  
qa.run(query)
```

"Zsuzska stole the devil's tenger-ütő pálczája (sea-beating stick), tenger-lépő cipője (sea-stepping shoes), and arany kis gyermek (golden baby) in an arany bölcső (golden cradle). She also previously stole the devil's tenger-ütőpálczát (sea-beating stick) and arany fej káposztát (golden head cabbage)."

Hiába tagadta, hogy nem volt hajlandó megelőzni a fiatal lányt, hiszen a fiatal lány mindenkit megtörte, amit fogott. De Zsuzska nem volt hajlandó megelőzni a fiatal lányt, hiszen a fiatal lány mindenkit megtörte, amit fogott.

De Zsuzska nem volt hajlandó megelőzni a fiatal lányt, hiszen a fiatal lány mindenkit megtörte, amit fogott.

– Nem, nem, nem, nem, nem,

Zsuzska azt is el tudta mondani,

lopja, halálnak halálával halmeg. Elindult hatodik napra,

arany fej káposztát, avval bekiáltott az ablakon. – Hej ördög, vissza-

megölettek három szép lyányomat, elloptad a tenger-lépő cipőmet, elloptad a tenger-

káposztám, csak ezt az egyet add vissza, soha szemedre sevetem.

fej
ák,
nek
ly
tudod

ne
is,
nem

vágta az
a,
ej

Can GPT still hallucinate? **Of course!**

Can I verify the answers? **Kind of!**

Does this allow you to easily "chat" with documents, even across languages??

Absolutely!!!!



and thus our questions and GPT
and the embeddings and the golden
cabbages all lived happily ever after

Jonathan Soma

Knight Chair in Data Journalism, Columbia University

js4571@columbia.edu

@dangerscarf

<https://github.com/jsoma/mediaparty-folktales>