After the release of ChatGPT in 2022, the number of papers published every day all pout Large Language Models (LLMs) has in creased more than 20-fold. The number of parameters in these LLMs jumped from of parameters in these LLMs jumped from the second content of parameters.	se of their c
of parameters in these LLMs jumped from of parameters liese LLN	
340 millions in implementations such as B ERT to billions of parameters in models like EPT-3 or LLaMA. A large part of these pa	of pari tese LEN
arameters come from the word-embeddir ig layers which are used to represent a fin ite vocabulary of characters, sets of chara	repr
del complexity, a fixed vocabulary is also re esponsible for brittle models, which canno t deal with out-of-vocabulary inputs and ca	m tames. Ap
After the release of ChatGPT in 2022, the number of papers published every day al	se of that (
oout Large Language Models (LLMs) has in creased more than 20-fold. The number of parameters in these LLMs jumped from 340 millions in implementations such as B mplementation	r c ta,000. The section of part the section of
ERT to billions of parameters in models like GPT-3 or LLaMA. A large part of these parameters come from the word-embedding	ike 30 like
ite vocabulary of characters, sets of characters or words. Apart from increasing model complexity, a fixed vocabulary is also read to the complexity of characters, sets of characters or words. Apart from increasing model complexity, a fixed vocabulary is also read to the complexity of characters, sets of characters or words. Apart from increasing model complexity of characters, sets of characters or words. Apart from increasing model complexity of characters, sets of characters or words. Apart from increasing model complexity of characters, sets of characters or words. Apart from increasing model complexity of characters, sets of characters or words. Apart from increasing model complexity or words.	cte temesis. Ap easing
esponsible for brittle models, which cannot esponsible for brittle models, which cannot deal with out-of-vocabulary inputs and cause with out-of-vocabulary inputs and	espon with but-c language
After the release of ChatGPT in 2022, the number of papers published every day all pout Large Language Models (LLMs) has in creased more than 20-fold. The number of the n	nt film wellstich ge educations
of parameters in these LLMs jumped from 340 millions in implementations such as B ERT to billions of parameters in models like T to billions of	of par nese in N 18 ntinibile less of like Tibile less of like
e GPT-3 or LLaMA. A large part of these part of the part of these part of the part of	dir are repr
cters or words. Apart from increasing models complexity, a fixed vocabulary is also reasonable for brittle models, which cannot tell with out-of-vocabulary inputs and causes a second words. Apart from increasing models are second words. Apart fr	re m sasing not assing
nnot generalize to new languages. As a language language. After the release of ChatGPT in 2022, the language language.	After the se of the C
number of papers published every day at pout Large Language Models (LLMs) has in a creased more than 20-fold. The number of parameters in these LLMs jumped from a proper publish and the property of parameters in these LLMs jumped from the property of parameters in the property day at the property of parameters in the property day at the property of parameters publish and property of parameters publish and property day at the property of parameters publish and propert	ge Lang Bar her control of par hese 121 N
340 millions in implementations such as B ERT to billions of parameters in models like GPT-3 or LLaMA. A large part of these parameters come from the word-embeddin	like TELLING HELDER HINE PARTY CONTROL
ig layers which are used to represent a fin ig layers which are used to represent a fin ite ocabulary of characters, sets of characters or words. Apart from increasing models to the original sets of the original sets or	te cabu te sain nox
del complexity, a fixed vocabulary is also responsible for brittle models, which cannot deal with out-of-vocabulary inputs and cannot generalize to new languages. As a second control of the second c	espon the same is a spon to
After the release of ChatGPT in 2022, the number of papers published every day al pout Large Language Models (LLMs) has in	After the se of fluit (
of parameters in these LLMs jumped from 340 millions in implementations such as B 340 millions in mplementation	of par hestern
ERT to billions of parameters in models like GPT-3 or LLaMA. A large part of these parameters come from the word-embedding layers which are used to represent a fingle layers which are used to represent a final layer	pa e A lar pa direct nur er of secret
ite vocabulary of characters, sets of characters or words. Apart from increasing model complexity, a fixed vocabulary is also reasonsible for brittle models, which cannot esponsible for brittle models.	re m artistic events in hich canno
t deal with out-of-vocabulary inputs and cannot generalize to new languages. As a self-the release of ChatGPT in 2022, the	ot gene languag
number of papers published every day all pout Large Language Models (LLMs) has in acreased more than 20-fold. The number of the language is a contract of the language in the language in the language is a contract of the language in the language in the language is a contract of the language in the language in the language is a contract of the language in the language in the language is a contract of the language in the language	nt file delist file as all ge Lang still still the second
of parameters in these LLMs jumped from 340 millions in implementations such as B ERT to billions of parameters in models like GPT-3 or LLaMA. A large part of these page 1. To billions of the 1. A large part of these page 1. A large part of these page 1. A large part of the	ike Figure 1 Figure 2 Figure
arameters come from the word-embedding layers which are used to represent a fing layers which are used to represent a fing layers which are used to represent a fing layers which are used to represent a fing layers which are used represent a fing layers which are used to represent a fing layers which are used	te hass y series of characters and the series of the serie
del complexity, a fixed vocabulary is also responsible for brittle models, which cannot deal with out-of-vocabulary inputs and cannot generalize to new languages. As a second of the complexity of fixed vocabulary is also responsible for brittle models, which cannot generalize to new languages. As a second of the complexity of fixed vocabulary is also responsible for brittle models, which cannot generalize to new languages. As a second of the complexity of fixed vocabulary is also responsible for brittle models, which cannot generalize to new languages.	no sepen substanting to the sepanting to
After the release of ChatGPT in 2022, the number of papers published every day at	al number of the material of the second of t
oout Large Language Models (LLMs) has in creased more than 20-fold. The number of parameters in these LLMs jumped from 340 millions in implementations such as B 340 millions in mplementation.	creased e tha delimination of part terms to the little of
ERT to billions of parameters in models like ERT to billions of these parameters on the word-embedding layers which are used to represent a fining layers.	pa state and state of the see particle of the
ite vocabulary of characters, sets of characters of characters or words. Apart from increasing model complexity, a fixed vocabulary is also read the complexity a fixed vocabulary is also read to the complexity of characters, and the characters of characters	re de la company
esponsible for brittle models, which cannot deal with out-of-vocabulary inputs and caunot generalize to new languages. As a ot gene with out-of-words and caunot generalize to new languages.	ca system to tareit ny terp stated ca et gene switches to the same of the same
After the release of ChatGPT in 2022, the number of papers published every day all pout Large Language Models (LLMs) has in creased more than 20-fold. The number of icreased nore than 20-fold.	
of parameters in these LLMs jumped from 340 millions in implementations such as B 10 millions of parameters in models like ERT to billions of parameters in models	om efective de la company de l
e GPT-3 or LLaMA. A large part of these part of these part of these part ameters come from the word-embedding layers which are used to represent a finite vocabulary of characters, sets of characters it evocabular of characters are the sets of these part of the par	ra
cters or words. Apart from increasing models complexity, a fixed vocabulary is also reasonsible for brittle models, which cannot be deal with out-of-vocabulary inputs and caute and caute of the complexity of th	no managada da
nnot generalize to new languages. As a nnot generalize new languages. A fter the release of ChatGPT in 2022, the number of papers published every day al	he
oout Large Language Models (LLMs) has in creased more than 20-fold. The number of parameters in these LLMs jumped from of parameters in these LLMs jumped from	om
340 millions in implementations such as B ERT to billions of parameters in models like ERT to billions of parameters in models like ERT to billions of parameters in models arameters come from the word-embedding arameters.	pa
ig layers which are used to represent a fin ite vocabulary of characters, sets of characters or words. Apart from increasing modulel complexity, a fixed vocabulary is also resided as a fixed vocabulary is also resided.	10(
esponsible for brittle models, which cannot deal with out-of-vocable ry inputs and cannot generalize to new languages. As a nnot generalize to new languages. As a	no en la companya de