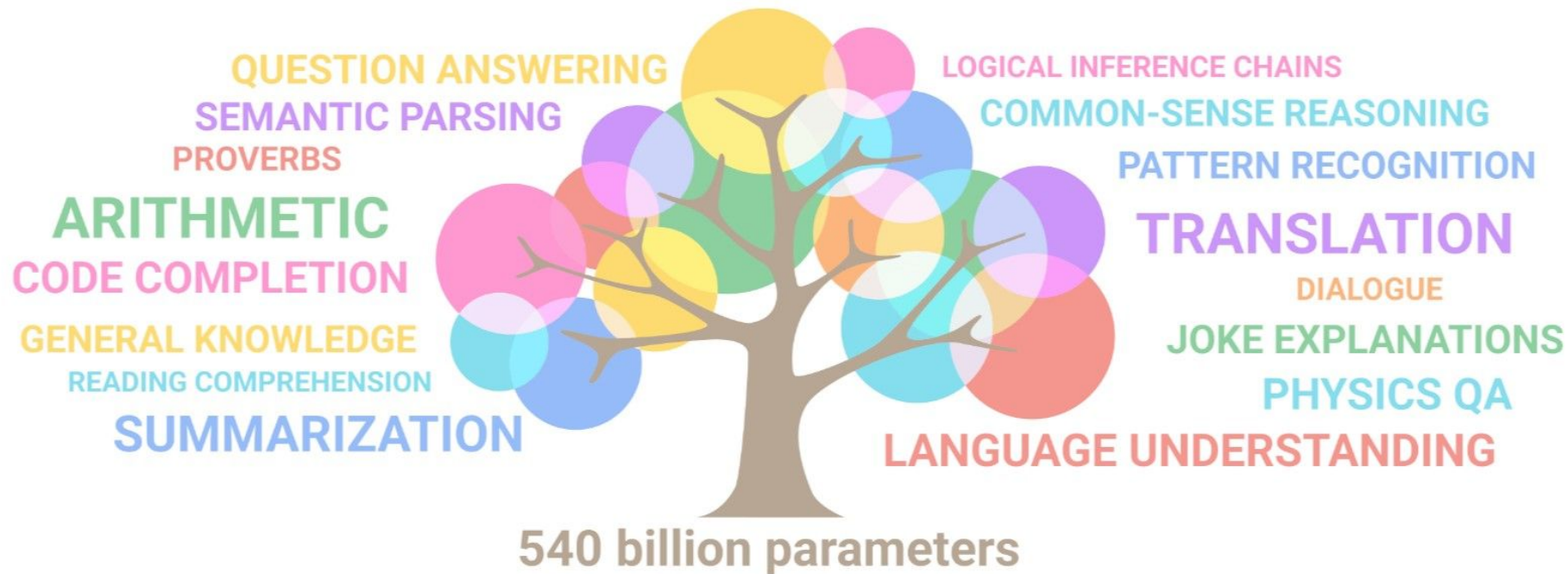


BLIP-2

Sara Godoy Franco - IIC3692

Modelos gigantes, capacidades increíbles



Más es mejor, pero ¿A que costo?



175B parameters
3.14E+23 FLOPS for training
355 GPU years
\$4.6M for a single training



11B parameters
3.3E+22 FLOPS for training
37 GPU years
\$480.000 for a single training



540B parameters
2.5E+24 FLOPS for training
2826 GPU years
\$36M for a single training

Visual-Language Research

— — —

Who is wearing glasses?

man



woman



Where is the child sitting?

fridge



arms



Is the umbrella upside down?

yes



no



How many children are in the bed?

2



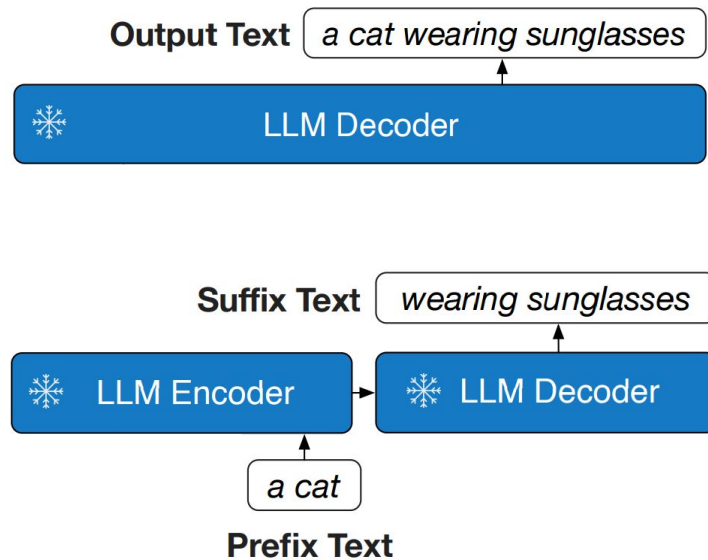
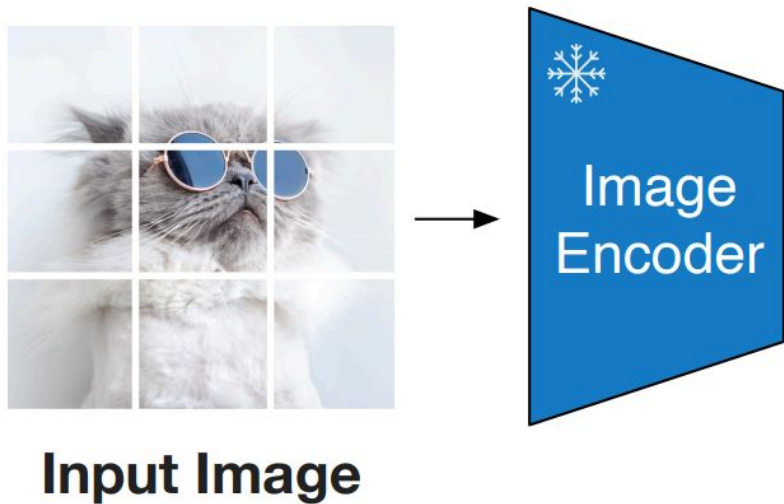
1



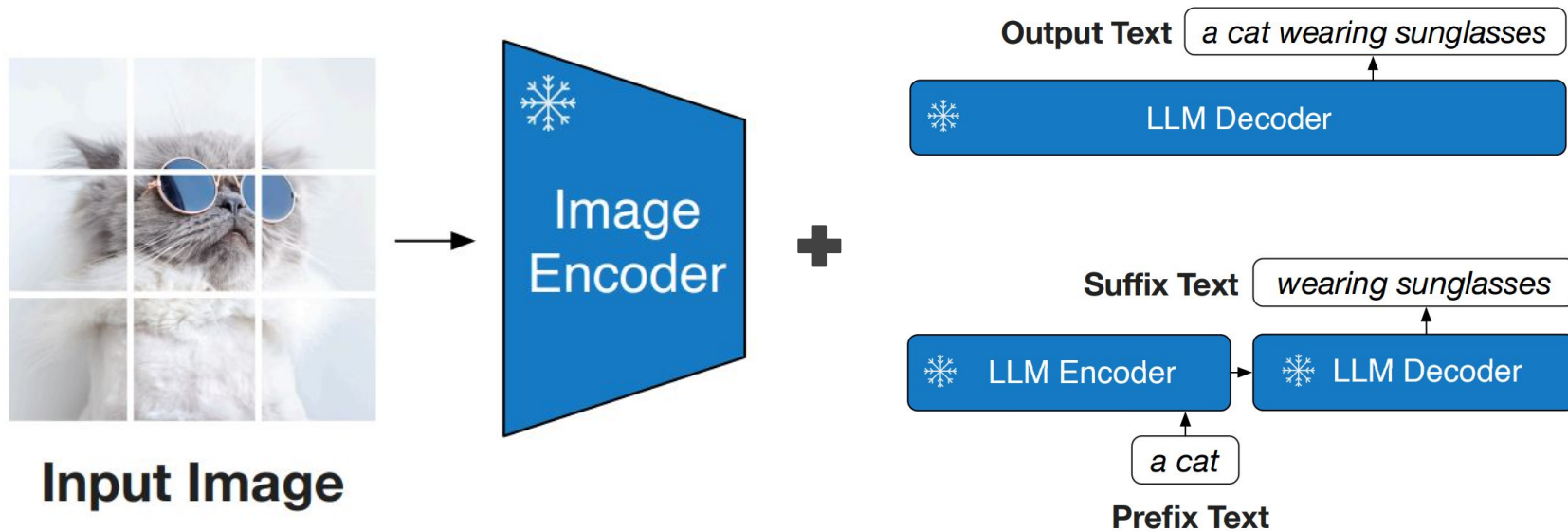
What color are her eyes?

What is the mustache made of?

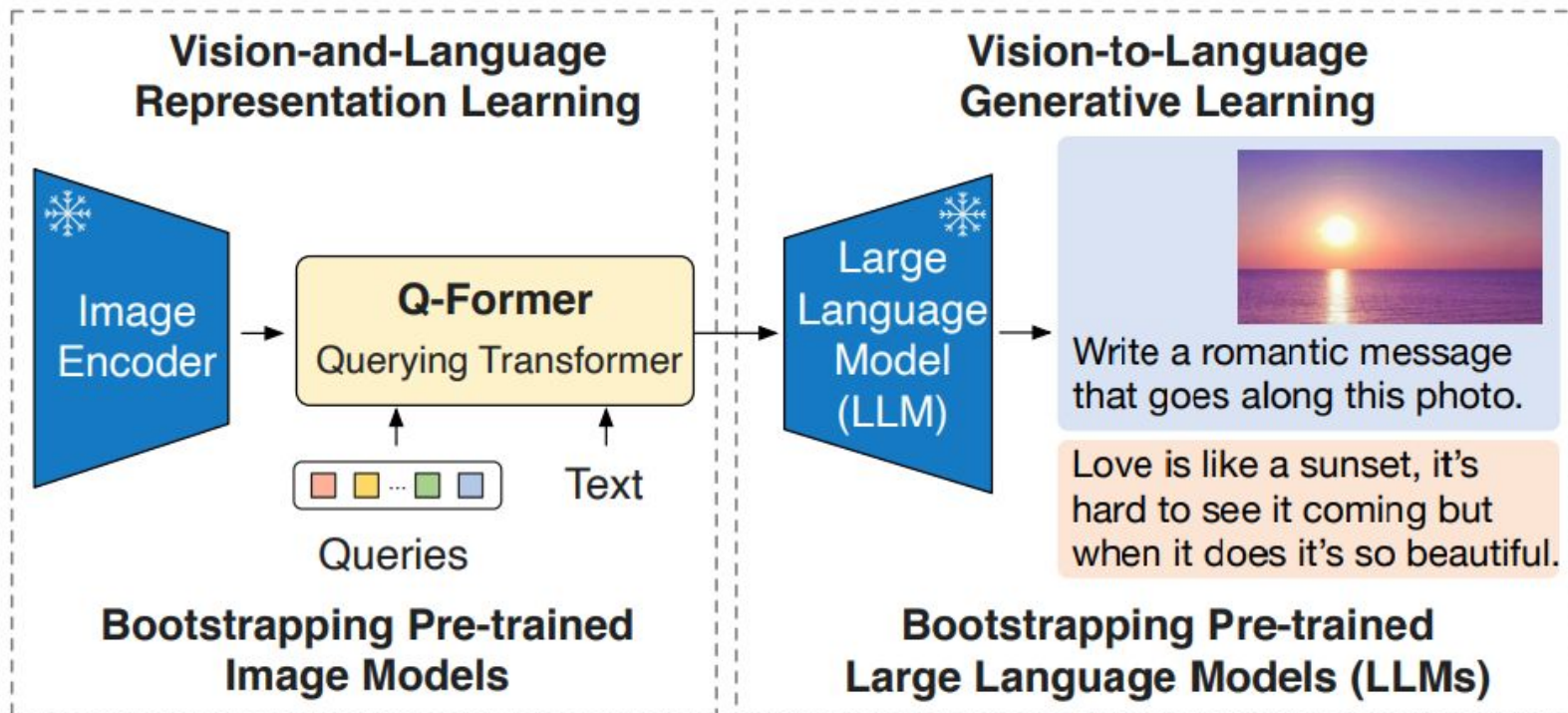
Juntando dos mundos



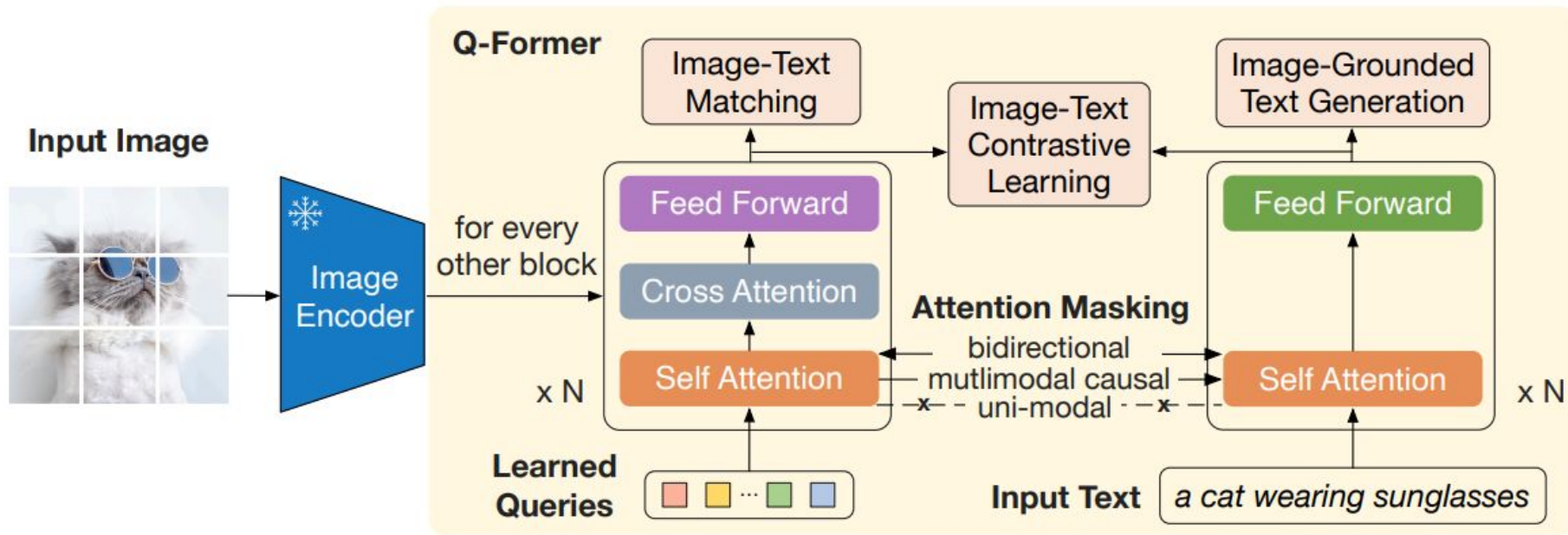
Bootstrapping Language-Image Pre-training



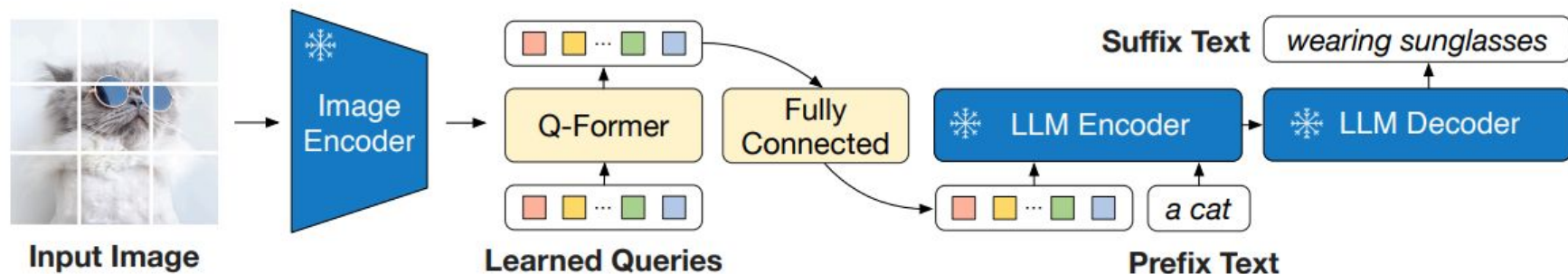
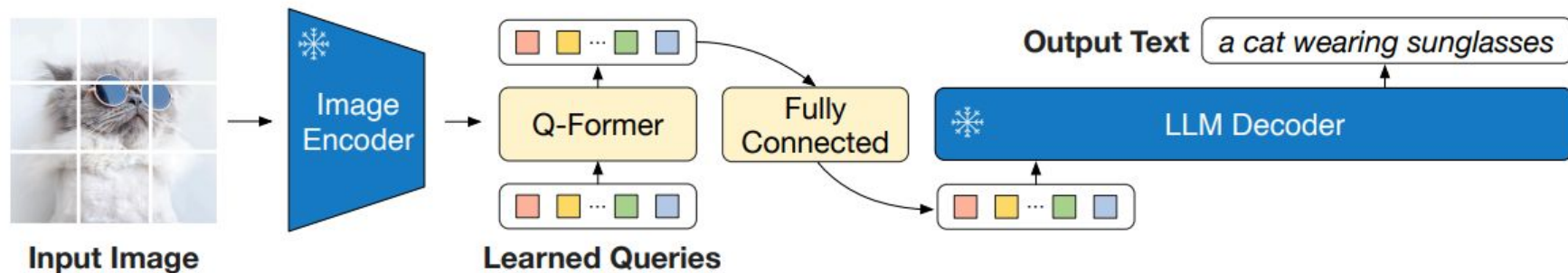
Q-Former : El puente entre dos mundos



First Stage : Visual-and-Language Representation Learning



Second Stage : Visual-to-Language Generative Learning





Explain the advantages of this product.

The audi e-tron quattro concept is a plug-in hybrid electric sports car that has a range of 310 miles and can accelerate from 0 to 60 mph in just 4 seconds.



Tell me something about the history of this place.

The great wall of china was built in 221 bc by emperor qin shi Huang to protect the emperor's capital from invaders from the north.



Write down the facts that you know about this flower.

It is a genus of flowering plants in the orchid family, orchidaceae. It is native to eastern asia and south east asia. It is a perennial flowering plant that grows up to a height of 1 m.



Is this photo unusual?

Yes, it's a house that looks like it's upside down.

How could someone get out of the house?

It has a slide on the side of the house.



What are shown in the photo?

A man and a chicken.

What does the man feel and why?

He is scared of the chicken because it is flying at him.



What are the ingredients I need to make this?

Pepperoni, mozzarella cheese, pizza sauce, olive oil, salt, pepper, basil.

What is the first step?

Place the pizza dough on a baking sheet, brush with olive oil, sprinkle with salt, pepper, and basil.



Quizás más no es siempre mejor

Models	#Trainable Params	#Total Params	VQAv2		OK-VQA	GQA
			val	test-dev	test	test-dev
VL-T5 _{no-vqa}	224M	269M	13.5	-	5.8	6.3
FewVLM (Jin et al., 2022)	740M	785M	47.7	-	16.5	29.3
Frozen (Tsimpoukelli et al., 2021)	40M	7.1B	29.6	-	5.9	-
VLKD (Dai et al., 2022)	406M	832M	42.6	44.5	13.3	-
Flamingo3B (Alayrac et al., 2022)	1.4B	3.2B	-	49.2	41.2	-
Flamingo9B (Alayrac et al., 2022)	1.8B	9.3B	-	51.8	44.7	-
Flamingo80B (Alayrac et al., 2022)	10.2B	80B	-	56.3	50.6	-
BLIP-2 ViT-L OPT _{2.7B}	104M	3.1B	50.1	49.7	30.2	33.9
BLIP-2 ViT-g OPT _{2.7B}	107M	3.8B	53.5	52.3	31.7	34.6
BLIP-2 ViT-g OPT _{6.7B}	108M	7.8B	54.3	52.6	36.4	36.4
BLIP-2 ViT-L FlanT5 _{XL}	103M	3.4B	62.6	62.3	39.4	<u>44.4</u>
BLIP-2 ViT-g FlanT5 _{XL}	107M	4.1B	<u>63.1</u>	<u>63.0</u>	40.7	44.2
BLIP-2 ViT-g FlanT5 _{XXL}	108M	12.1B	65.2	65.0	<u>45.9</u>	44.7