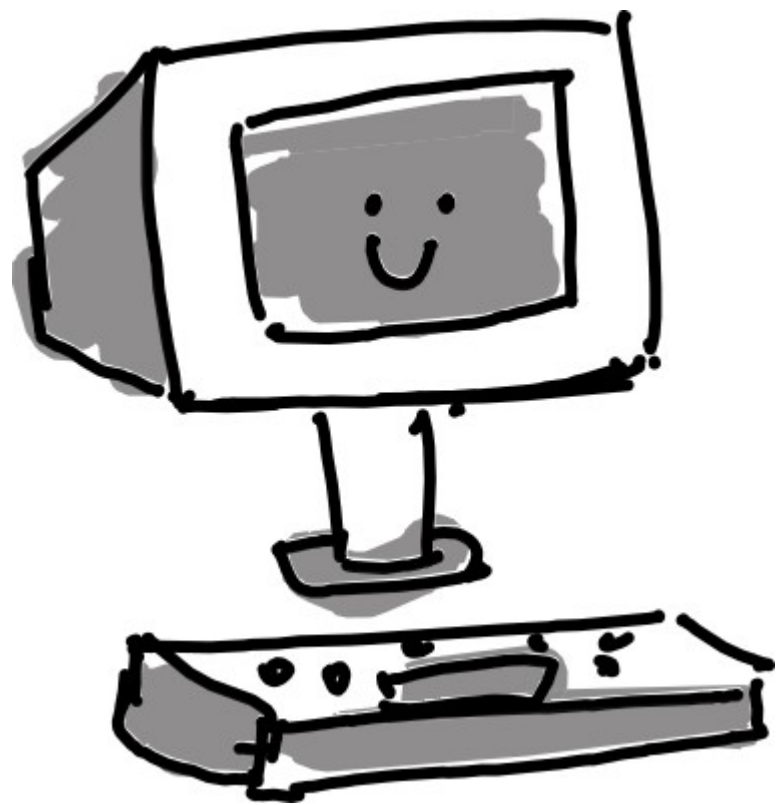
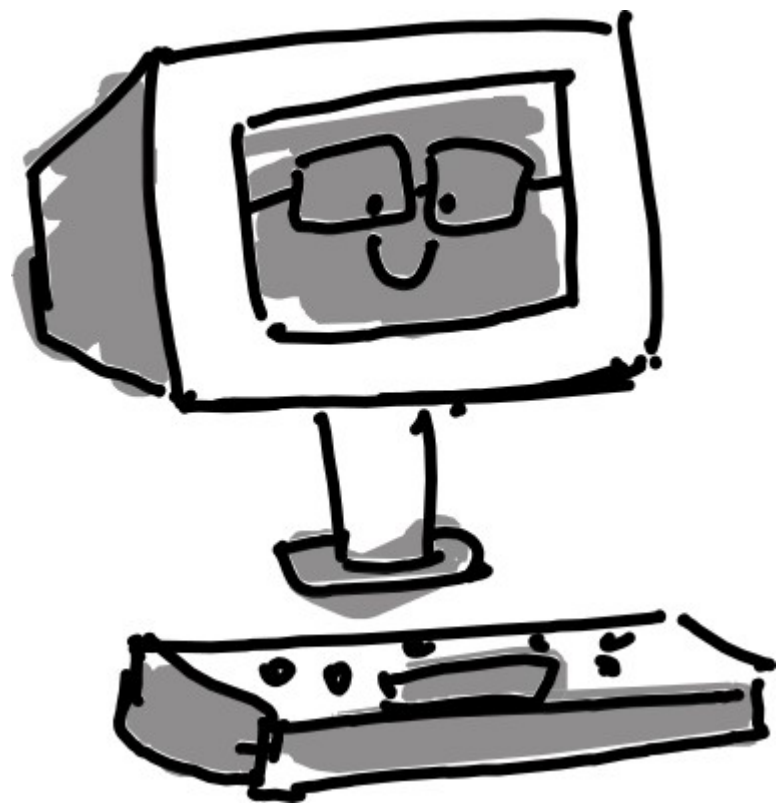


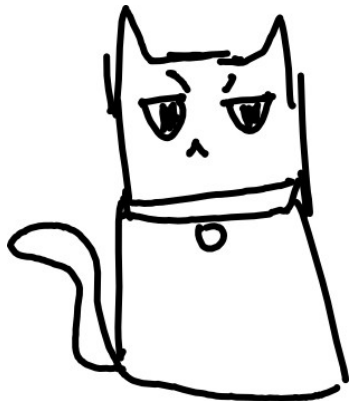
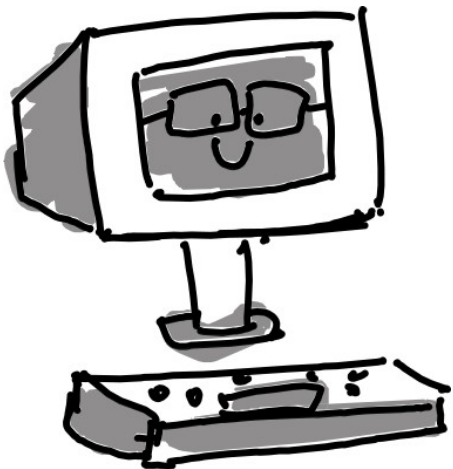
Large Audio Language Models  
are like cats: they listen to  
you only to ignore you

Giovana Morais  
May 1st  
Queer Science on Tap

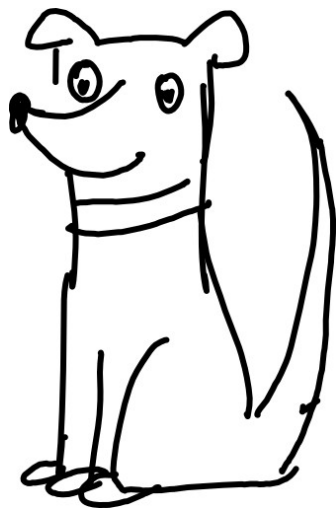






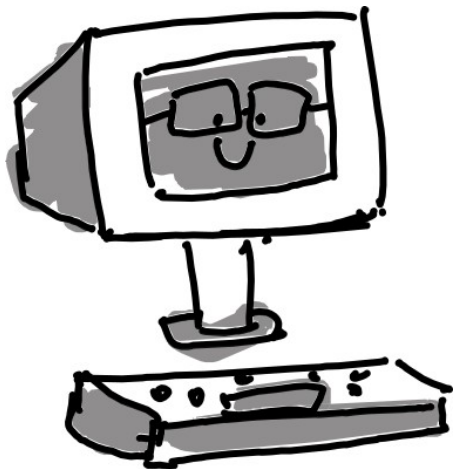


cat



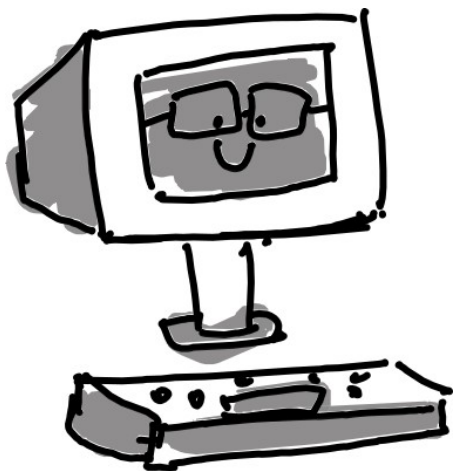
dog

cat!



(yes, this is my cat)



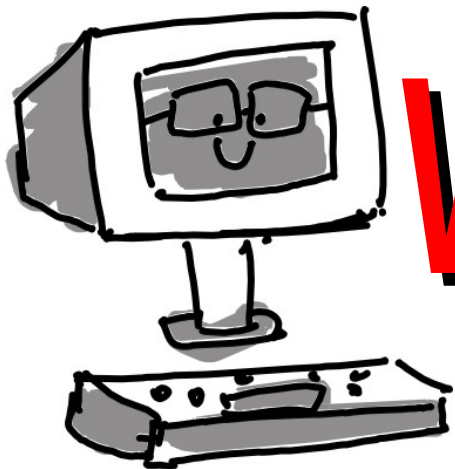


wolf



husky

wol̂f!

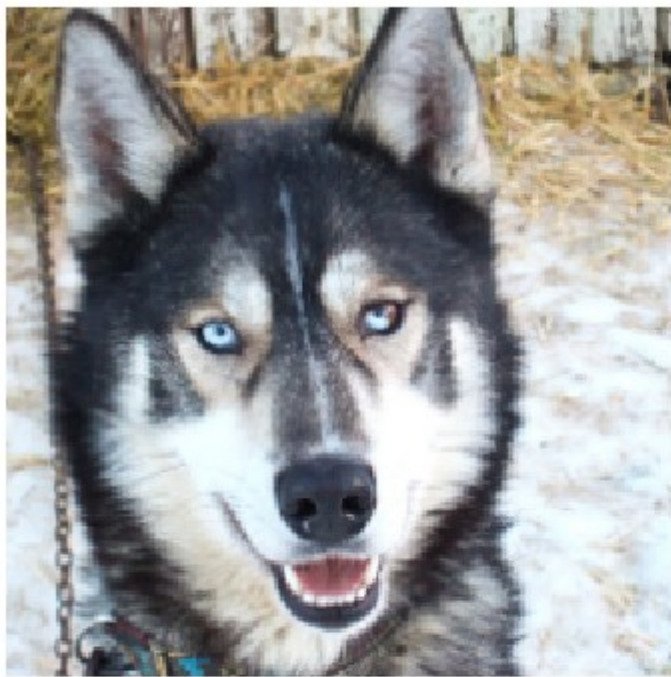


**WRONG!**

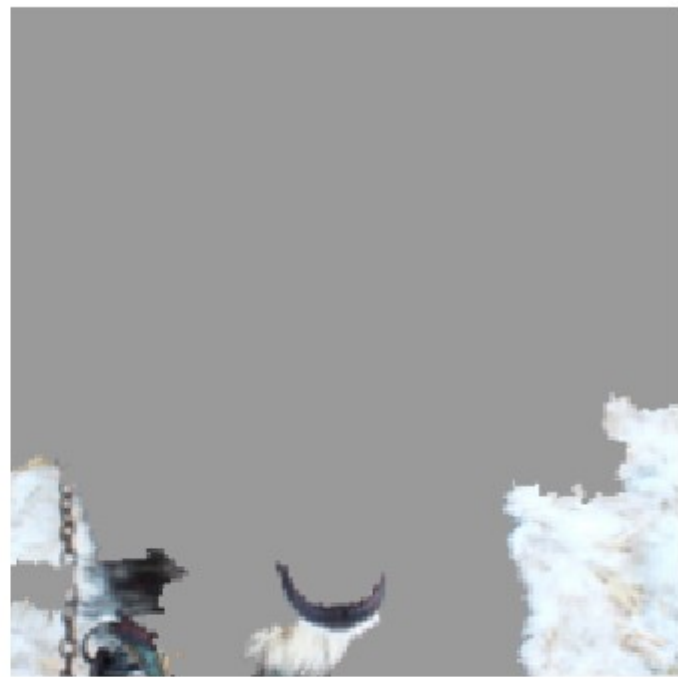




what is the  
computer  
paying  
attention to?



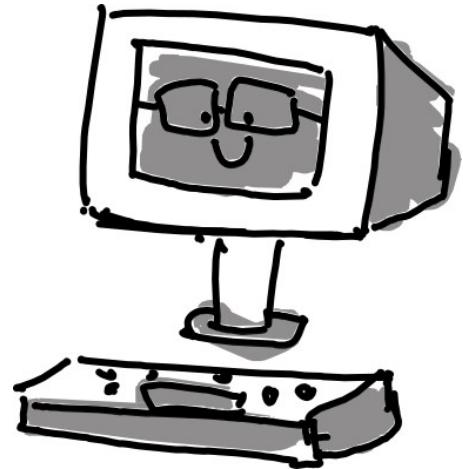
(a) Husky classified as wolf



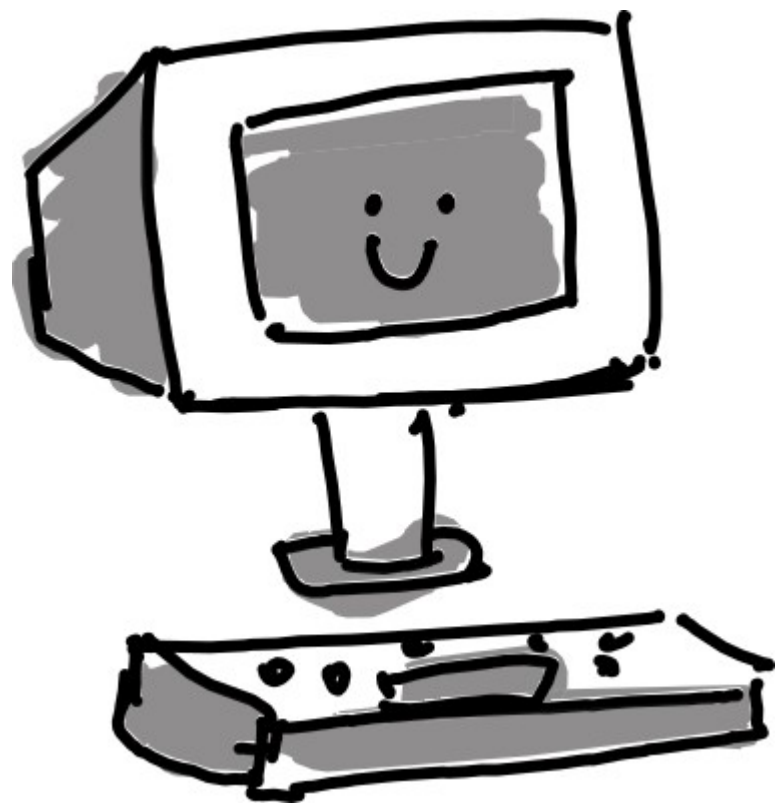
(b) Explanation

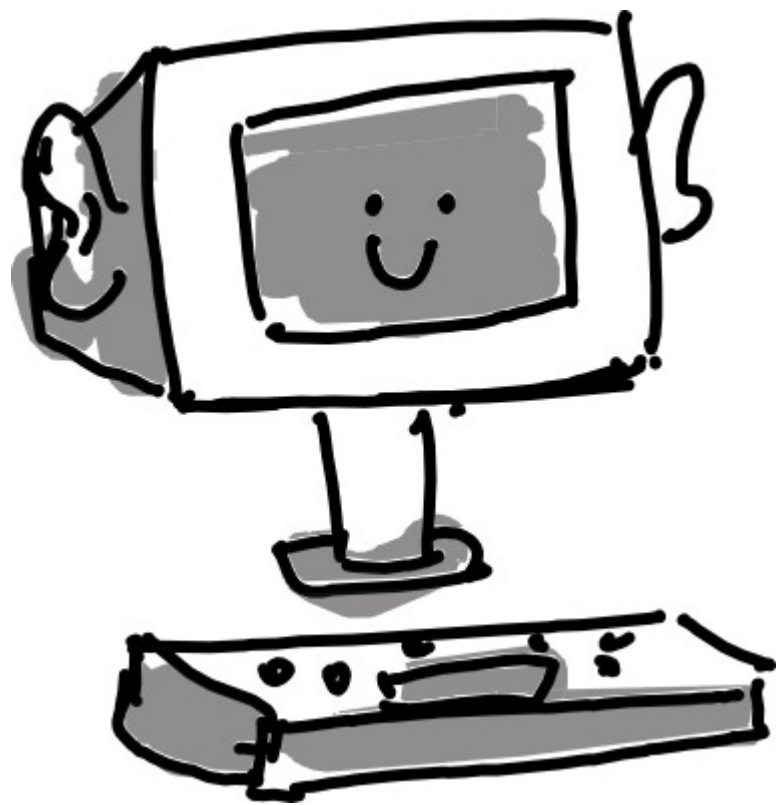
Computers are lazy  
and they will always  
look for shortcuts!

hehe, oops



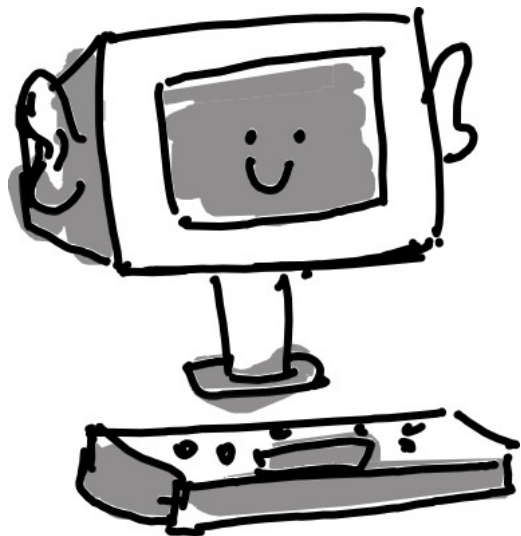
2026







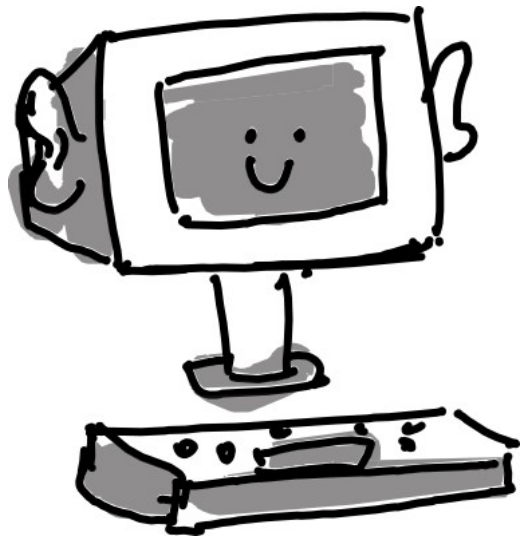
What is this  
instrument?



The instrument is a  
pandeiro.



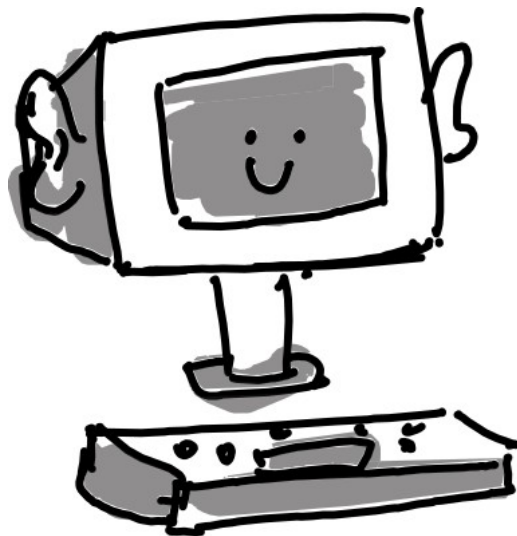
Give me the  
chord  
progression of  
the acoustic  
guitar



The instrument is a  
pandeiro.



What genre of  
music is this?

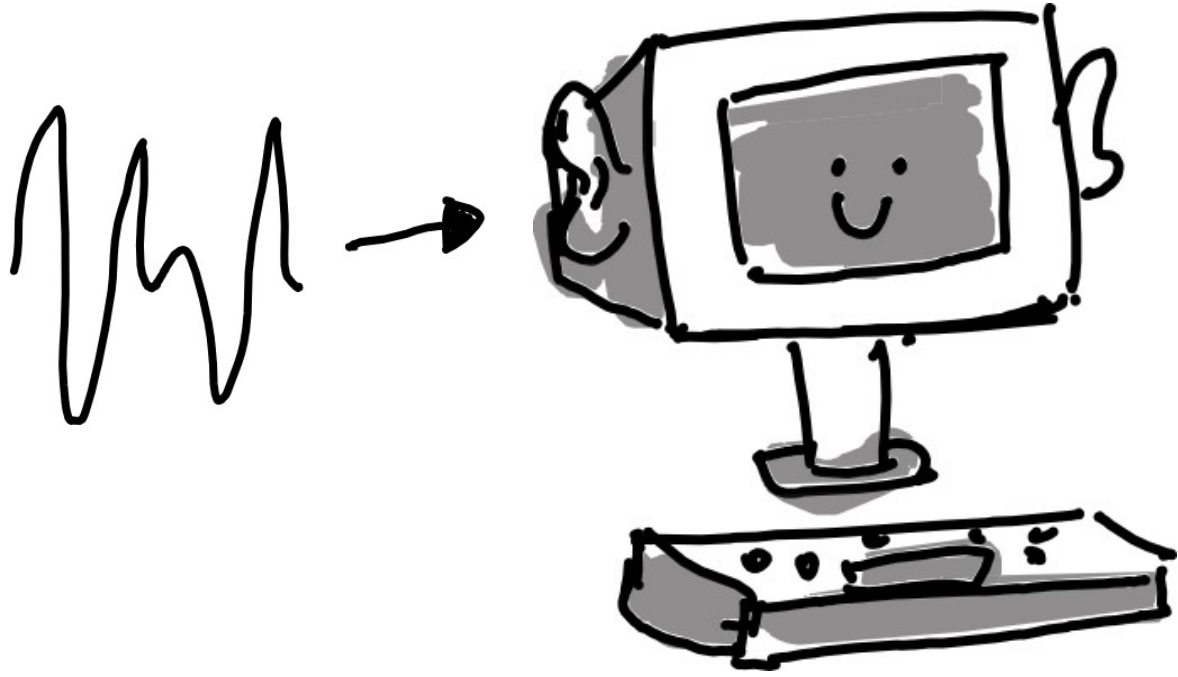


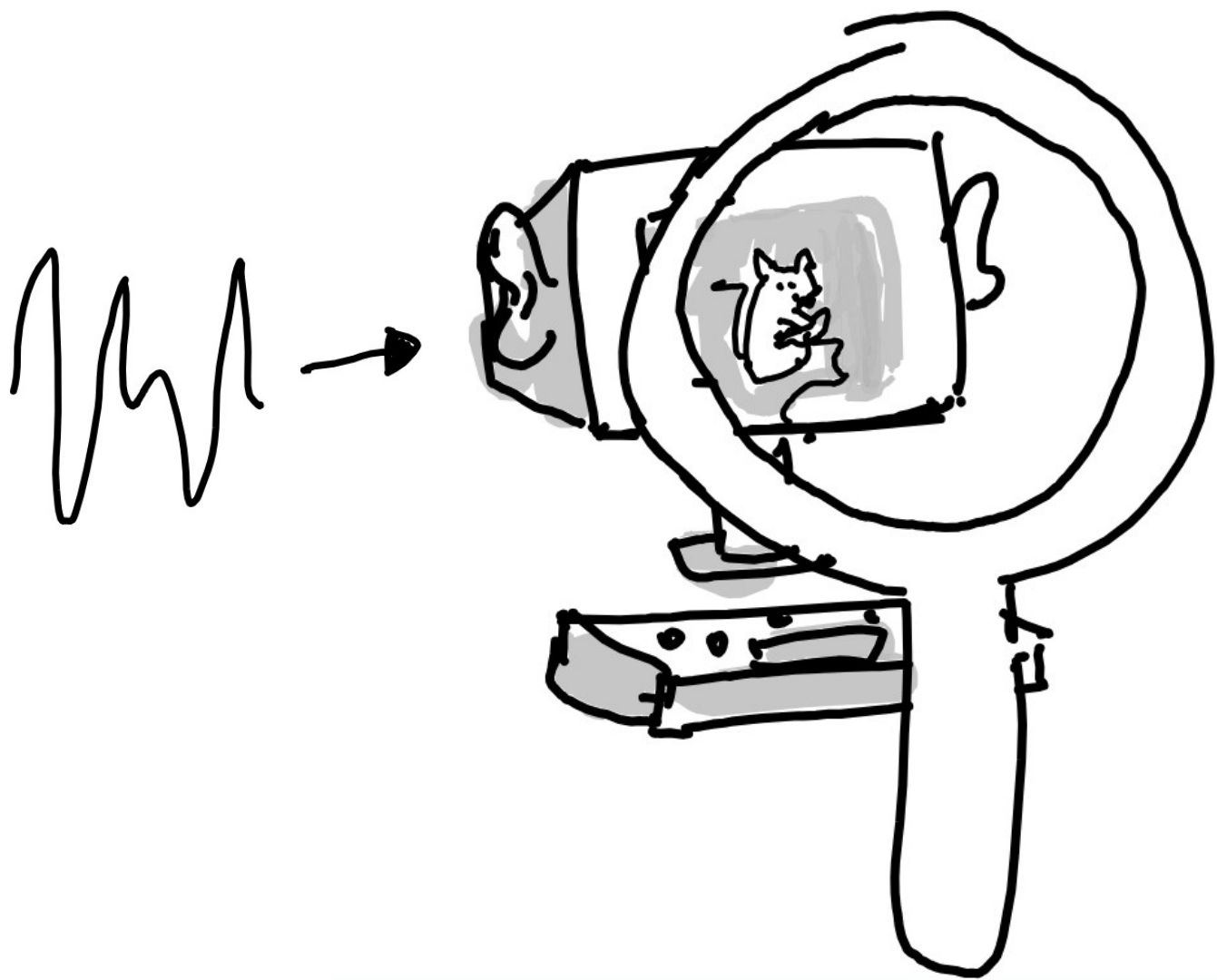
The instrument is a  
pandeiro.

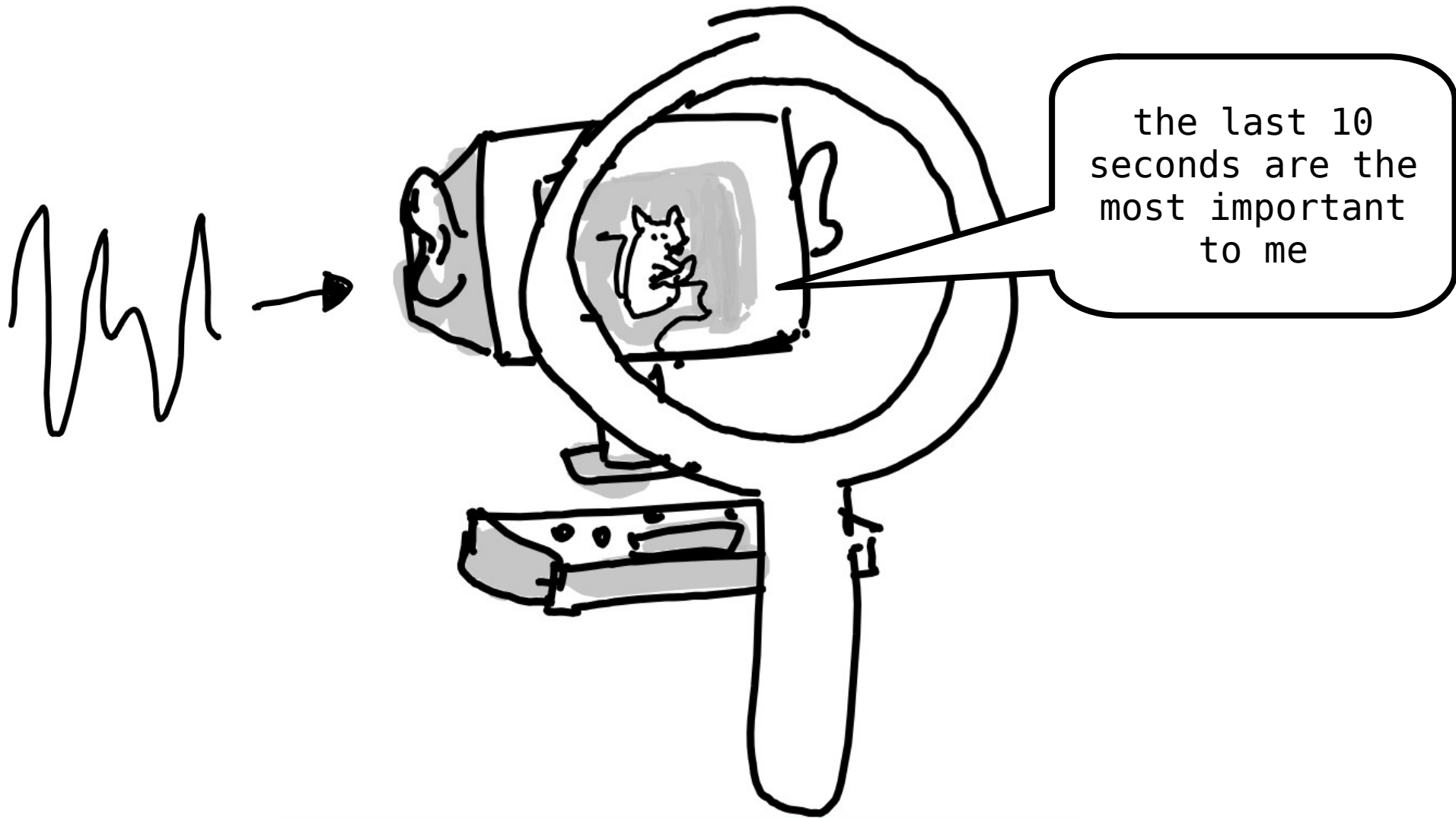


what is the  
computer  
paying  
attention to?

is it even  
listening to  
the audio?







the last 10  
seconds are the  
most important  
to me

# why is this important?

- transparency: we know what information is used
- trust: we agree that this information is the correct one to use, therefore we can trust the model
- reliability: if we trust the model, we can safely use it

so... are models listening  
to the audio?

yes!

but the answer is longer, more nuanced  
and more complicated than that.

(but also if you're curious **AND** willing to listen to me talking  
for more than 6 minutes, let's chat!)

obrigada!