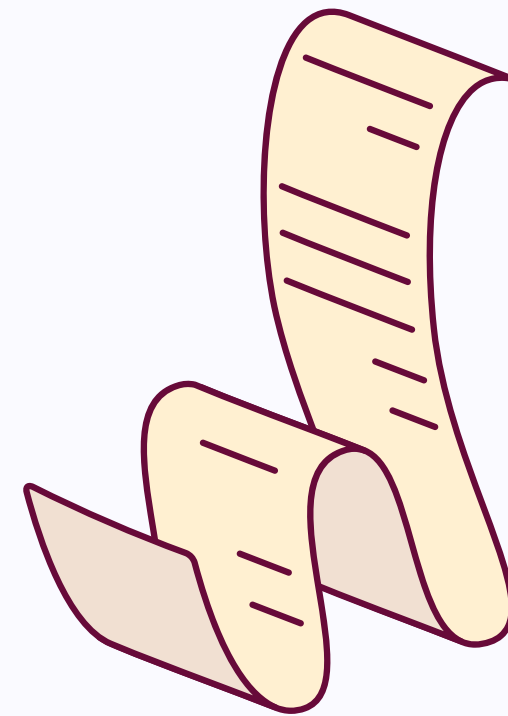


TEXT SUMMARIZATION WITH PYTHON

a gentle intro to natural
language processing



Ganning Xu
xu23g@ncssm.edu



Main Points

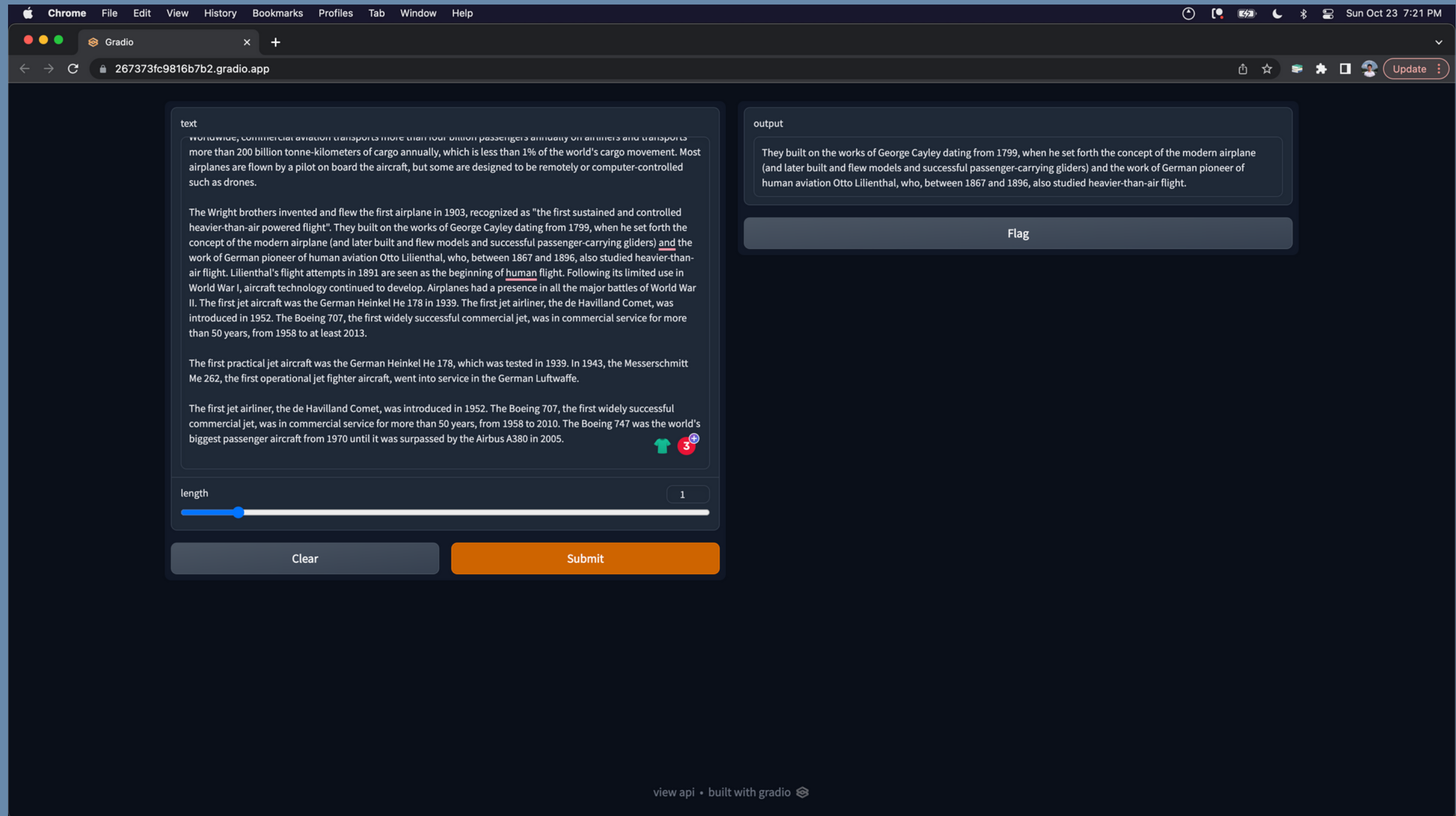
- Summary point 1
- Summary point 2
- Summary point 3

Hello!

- Senior at North Carolina School of Science and Mathematics
- Web Development & Machine Learning
- Building websites for nonprofit
- Creating random apps!
- Swimming, biking, airplanes



Final Product



What we're doing today:

1

Intro to natural language processing?

2

Types of text summarization

3

Commonly used for text summarization

4

Let's build out the text summarizer!

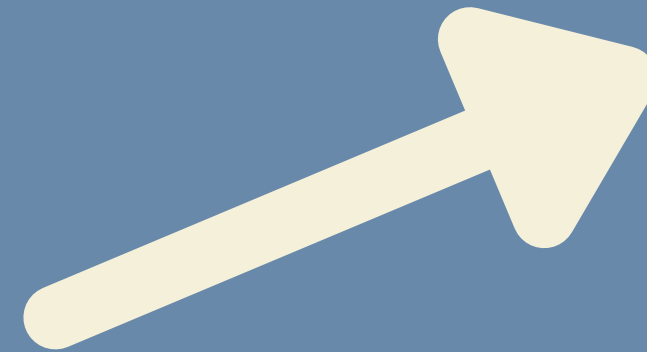
5

Deploy the model using Gradio



Imagine a world where...

Bob: This movie was absolutely terrible, I hated every moment of it. I almost fell asleep from how boring it was.



Bad movie

When computers understand human words:



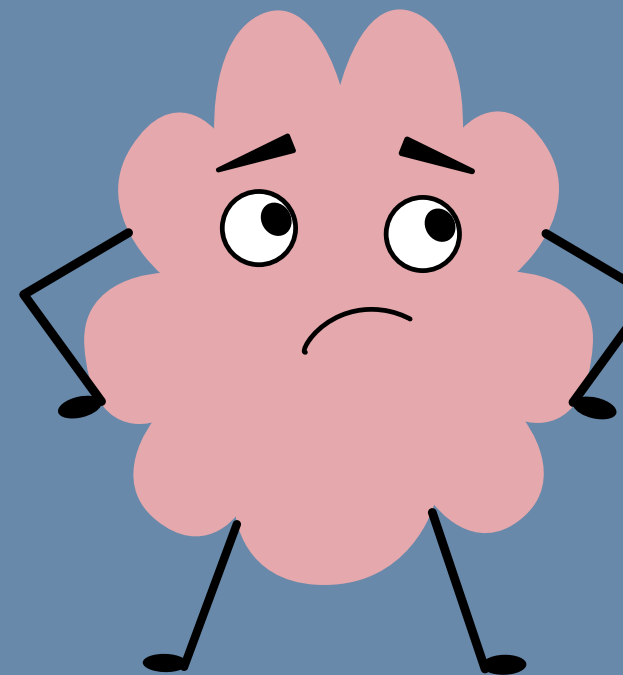
**Natural
Language
Processing!**

Why do we care?

- Personal assistants (siri, alexa)
- Flagging spam posts on Twitter
- Speech recognition: speech → text
- Text generation: writing your English essays
- and of course... **text summarization**

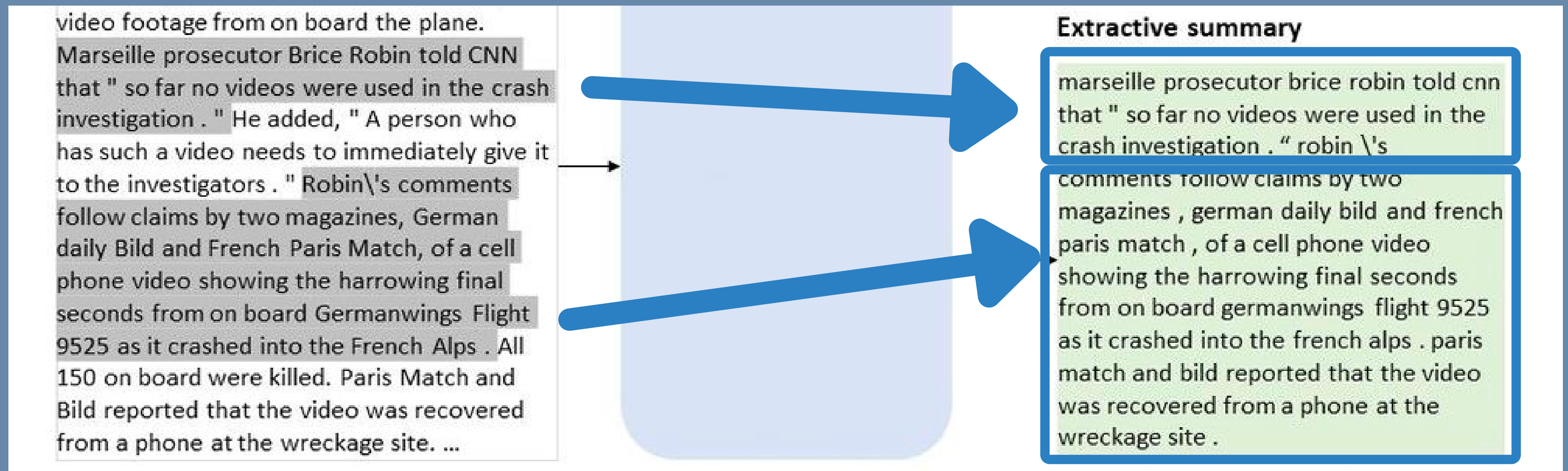
So, how do we do this?

Imagine if you were asked to
summarize a piece of text...
what would you do?



Cut out sentences with repeated info

extractive text summarization



Understanding the text, and making your own summary

abstractive text summarization

Input Article

Marseille, France (CNN) The French prosecutor leading an investigation into the crash of Germanwings Flight 9525 insisted Wednesday that he was not aware of any video footage from on board the plane. Marseille prosecutor Brice Robin told CNN that "so far no videos were used in the crash investigation." He added, "A person who has such a video needs to immediately give it to the investigators." Robin's comments follow claims by two magazines, German daily Bild and French Paris Match, of a cell phone video showing the harrowing final seconds from on board Germanwings Flight 9525 as it crashed into the French Alps. All 150 on board were killed. Paris Match and Bild reported that the video was recovered from a phone at the wreckage site. ...

Generated summary

Prosecutor : " So far no videos were used in the crash investigation "

Which one do you think is better?

extractive vs. abstractive

Well, it's complicated :)

Extractive

- Easier to code
- Fast at summarizing
- All text will be readable!
- Summaries are similar to text
- Summary may leave out context (pronouns, references)

Abstractive

- VERY HARD to code
- Takes days to train
- Summaries are original!
- May produce non-readable summaries
- Incorrect summaries
- Very cool :)

We're doing:

extractive

summarization!