AI & Filmmaking

Nicola Cortesi

INTELLIGENCE AS A COMMODITY

We are shifting from a society based on human labour and capital to a society based on data and intelligence.

Intelligence is becoming a commodity, as electricity or water or gas, that arrive silently in our houses.

THE COST OF MOST PRODUCTS AND SERVICES WILL GO TO ZERO

Software development

RIP CODING
AI IS HERE

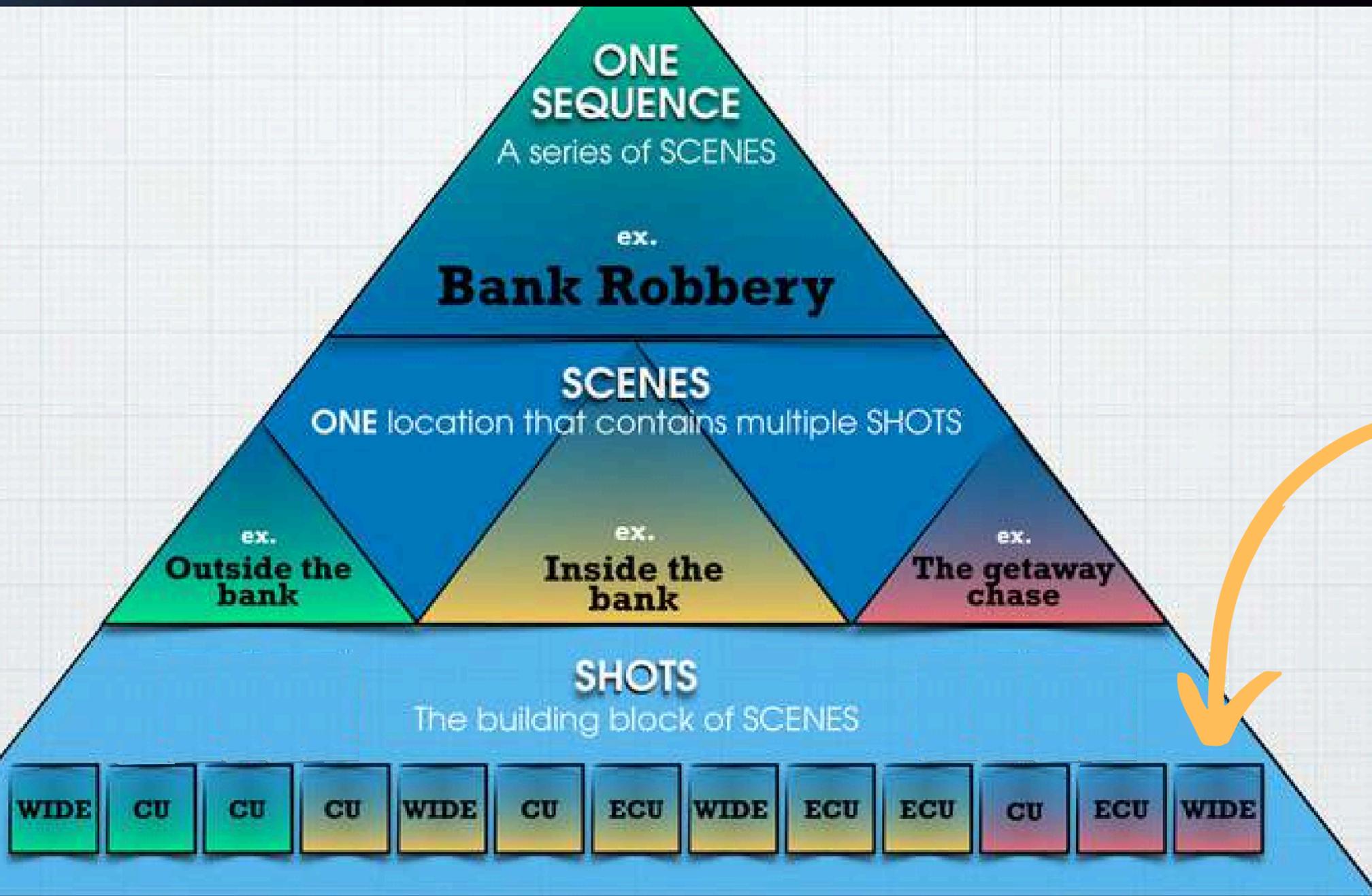**White Collar Jobs**

Open Claw
AI Agent

Manifacture

Dark factory

Energy

Solar farm

THE BASIC OF FILMMAKING: SEQUENCES, SCENES AND SHOTS

Shot: a continuous series of frames beginning from the moment the camera starts recording until it stops (in case of Google Veo, each shot corresponds to an 8-second video that Veo generates)

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Establishing Shot: a wide camera shot at the beginning of the movie or of a new scene, to inform the viewer of the location and time

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Master Shot: it usually follows the Establishing Shot, to show which characters are in the scene

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Wide Shot: Shows the entire subject from head to toe, to highlight his/her relationship with the location/environment

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Full Shot: Shows the entire subject, but only to capture the subject's physical identity and movements. There is very little space (headroom/footroom) above or below the subject.

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Medium Full Shot (Cowboy Shot): frames a subject from the mid-thighs up to the head. it suggests the audience that the character is about to do something—draw a weapon, grab a tool, or start a fight.

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Medium Shot: Frames the subject from the waist up. This is the most common shot in the cinema, particularly for dialogues.
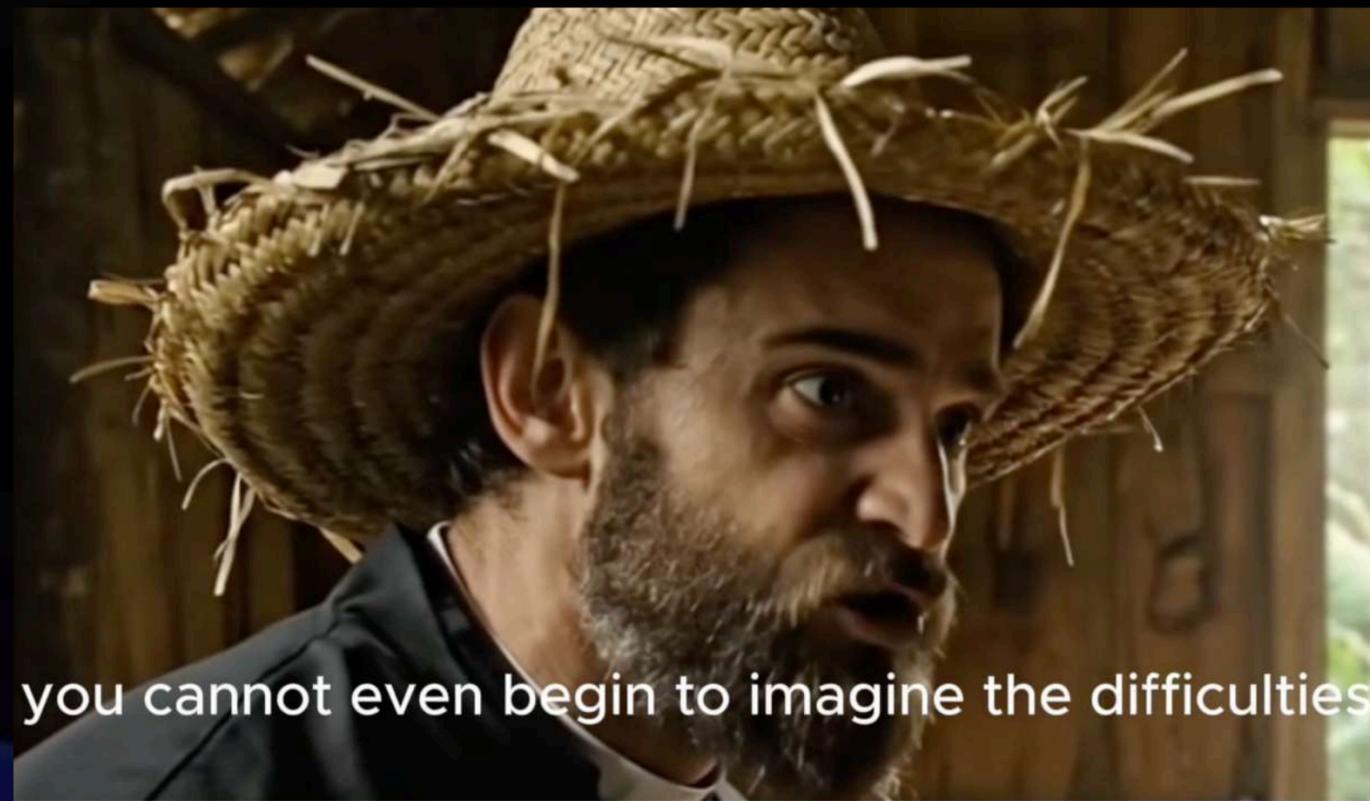
Camera Shots ("Shots"):

- **Medium close-up (MCU):** frames a subject from the chest or shoulders to the top of the head. A good compromise to get intimate with a subject without loosing its physicality

Camera Shots ("Shots"):

- Close-up (CU): Very common shot that ocuses on a specific part of the subject, usually the face, to convey emotions and empathy. Camera is usually at eye level

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- **Extreme Close-up (ECU):** It zooms in on a specific detail, like an eye or a ringing phone, for an intense emphasis.

Camera Shots ("Shots"):

- **Insert Shot:** close-up of a specific detail within a scene that highlights an object or a small action that might otherwise be missed. While most shots focus on characters and their performances, the Insert Shot focuses on information: a ticking clock, a text message on a phone, a hand pulling a trigger, or a key turning in a lock..

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Two Shot: a Two Shot is a camera framing where two subjects appear together in the same frame. It is the bread and butter of visual storytelling because it focuses on the relationship between characters rather than just the characters themselves. They are usually Medium Shots or Wide Shots or Close-ups

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Over the Shoulder (OTS): A shot taken from behind one person looking at another, used to establish a connection during conversations. It also create a sense of spatial depth

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Dirty Shot: a piece of an object or person in the foreground or in the background is intentionally left partially in the frame, but out of focus. It is called "dirty" because the subject's view is slightly "obstructed" or "cluttered" by something else, rather than being a "clean," unobstructed shot of the main subject. it makes the audience feel like they are spying on the scene.

SINGLE (DIRTY)

OTS (DIRTY)

A dirty two shot.

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Point of View (POV): Shows what a specific character is looking at, making the audience feel like they are in the character's shoes

THE GRAMMAR OF VISUAL STORYTELLING

Camera Shots ("Shots"):

- Extreme Wide Shot or Long Shot: The subject is very small, sometimes just a speck or a silhouette. Landscape (mountains, city, desert) takes up 80-90% of the frame. Goal: To emphasize scale, distance, isolation, or the epic nature of a journey (e.g., a lone traveler crossing a massive desert)

THE GRAMMAR OF VISUAL STORYTELLING

Camera Angles:

- **Eye level:** it is the most common camera angle, to follow the subjects. Dfault angle for dialoges.

THE GRAMMAR OF VISUAL STORYTELLING

Camera Angles:

- High Angle: the camera looks down, making the subject appear small, vulnerable, or powerless.

Camera Angles:

- **Low Angle:** the camera looks up at the subject, making them appear powerful, heroic, or intimidating.

THE GRAMMAR OF VISUAL STORYTELLING

Camera Angles:

- Ground Level: The camera is placed on the floor, often used to show people walking or to create a unique "bug's eye" perspective

Camera Angles:

- **Overhead / Bird's Eye View:** Shot from directly above, 90 degrees above the subject. It is often used to show layout or a sense of detachment.

Camera Angles:

- Dutch Angle: when the camera is tilted to a diagonal. Used to signal that something is "off" or wrong in the story. Used to create feelings of uneasiness, disorientation, madness, tension, or chaos.

Camera Angles:

- Aerial / Drone Shot: a shot taken from a high altitude (usually via drone or helicopter) to show vast landscapes.

measuring time had always been my obsession

THE GRAMMAR OF VISUAL STORYTELLING

Camera Angles:

- **Extreme wide angle:** capture both close-ups and landscapes in one shot. It makes the subject look tiny and insignificant compared to the vastness of the environment. It emphasizes isolation, loneliness, or just the epic scale of nature

EXTREME WIDE ANGLE

THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Pan (Panorama): The camera turns horizontally from left to right or right to left (like someone shaking their head to say "no"). It's used to follow a moving subject or to reveal more of a landscape

as if my very existence depended on it

Camera Motions:

- Tilt: the camera moves vertically up or down (like someone nodding "yes"). This is often used to show the scale of a tall building or to introduce a character from the feet up.

THE GRAMMAR OF VISUAL STORYTELLING
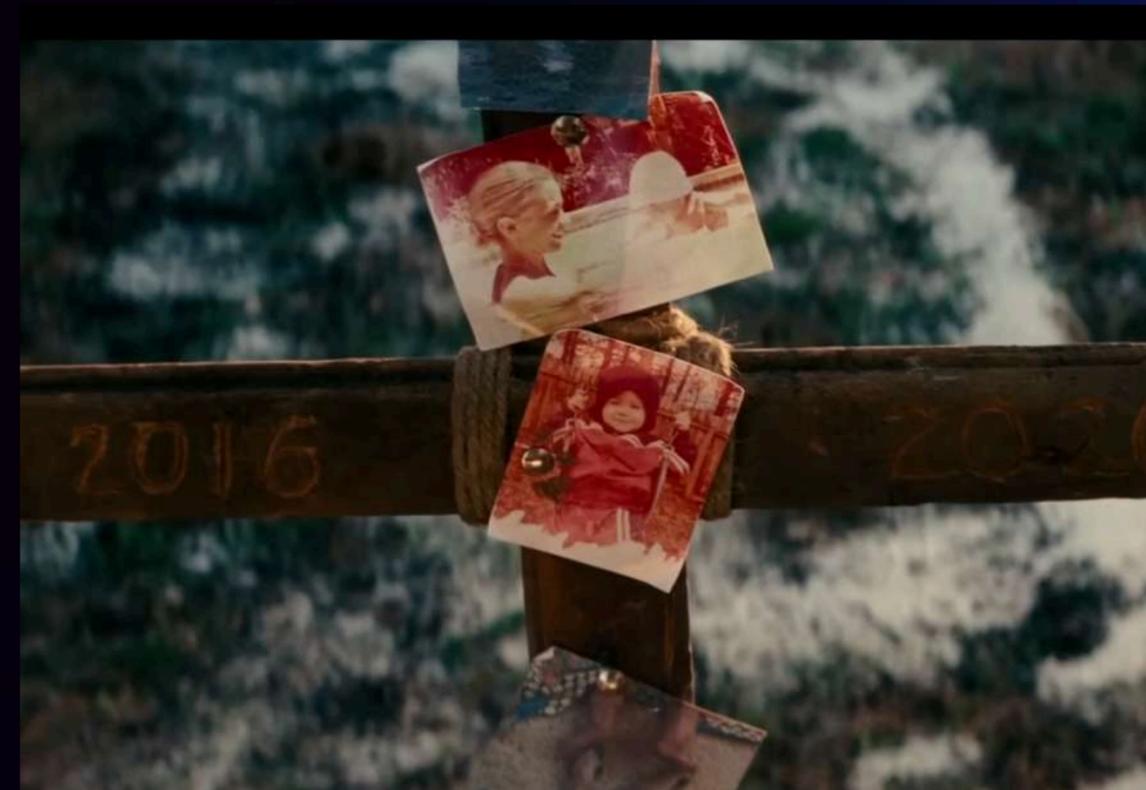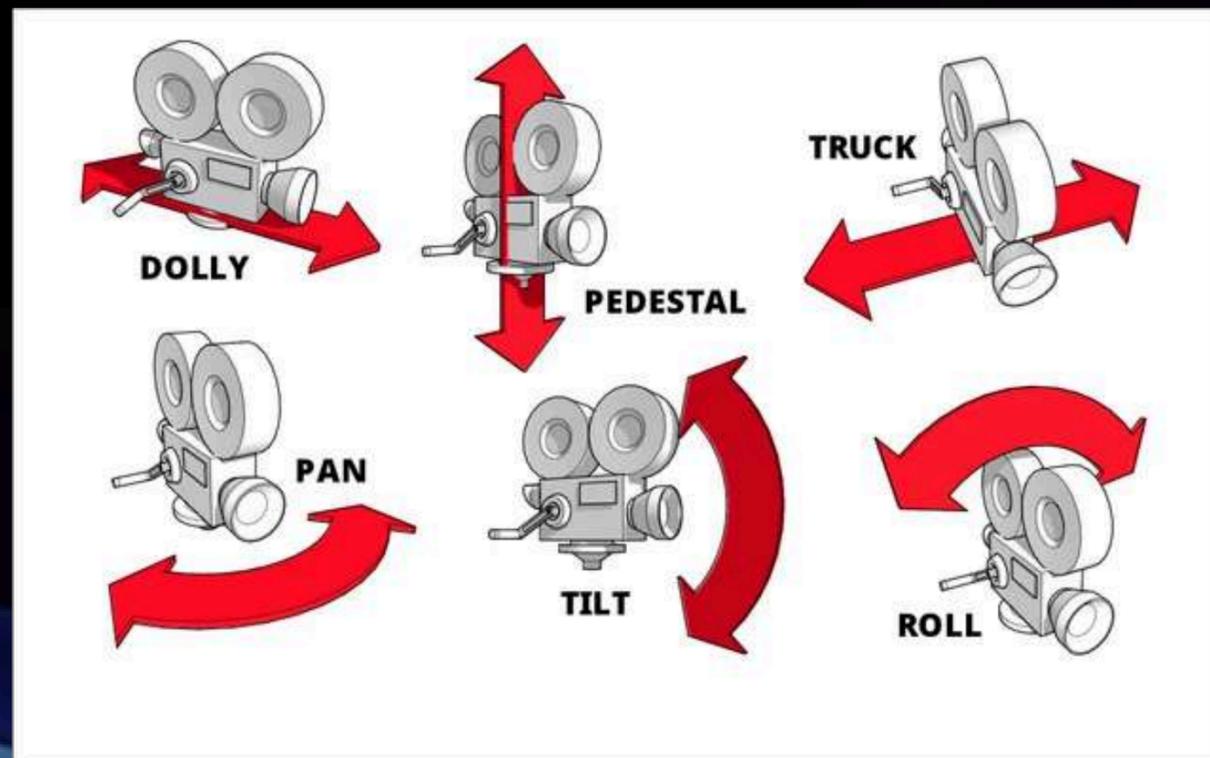
Camera Motions:

- Ped (Pedestal): moving the entire camera up or down without changing its vertical tilt. Imagine a camera on an elevator. Used to mantain eye level when a character sits down or stands up, or to build suspence
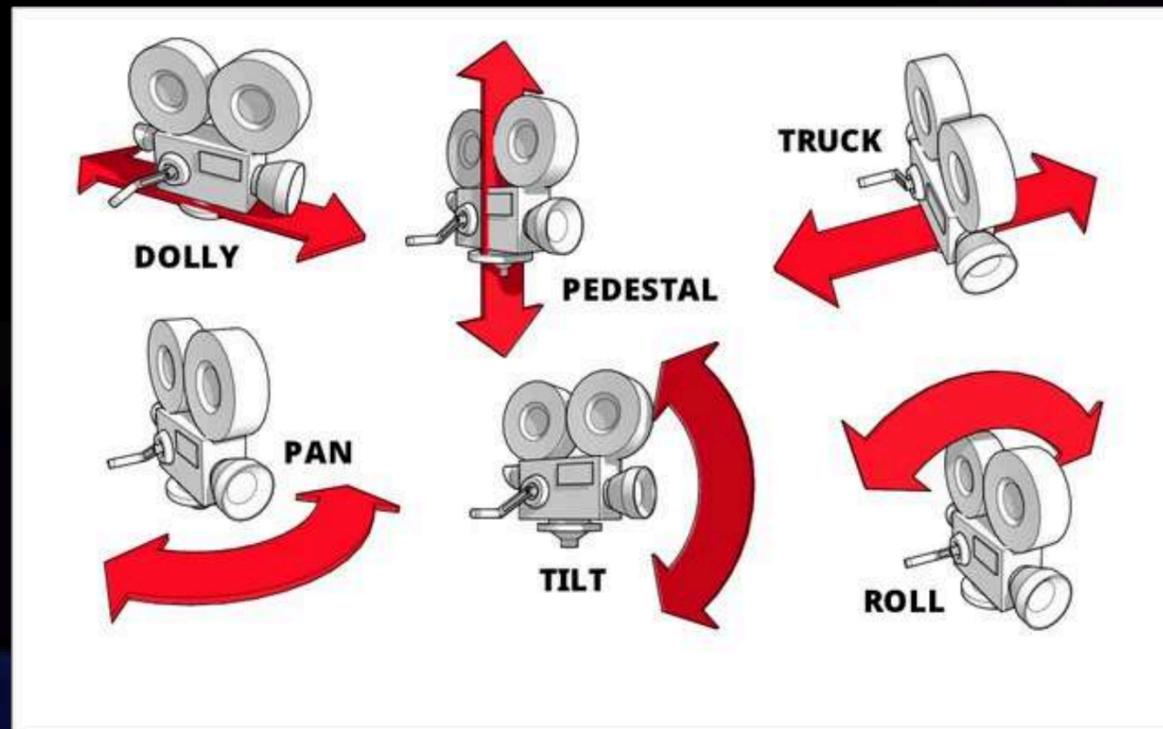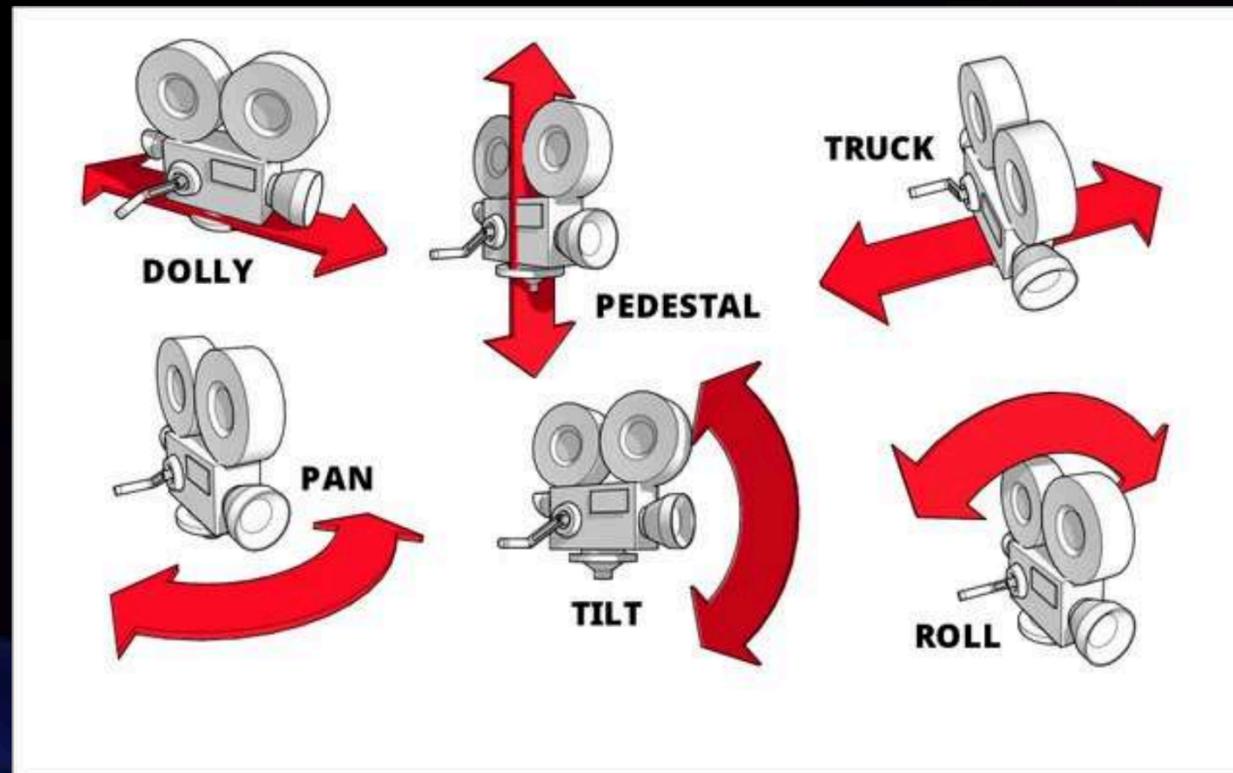
THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Dolly In/Push In: the camera is mounted on a track or wheels and moves toward the subject. It tells the audience that the moment is important, conveys internal conflict, or direct our attention to a specific detail. It feels more natural and immersive than a simple optical zoom in, that creates a sense of uneasiness. Use it instead of the zoom in!
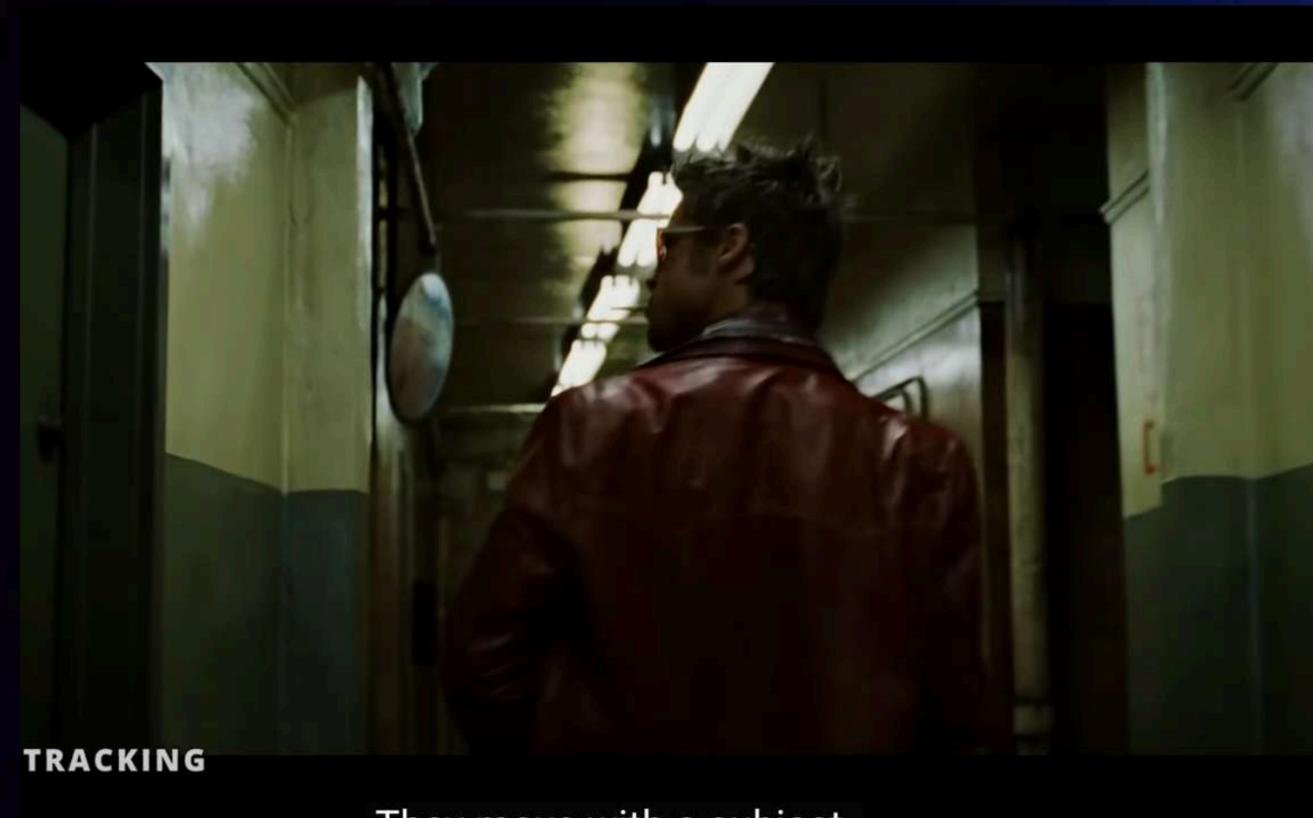
THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Dolly Out/Pull Out: the camera is mounted on a track or wheels and moves away from the subject. It unveils the context of a scene, disconnect from a character or from the whole film (last scene). It feels more natural and immersive than a simple optical zoom out. Use it instead of the zoom out!

THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Tracking Shot: A broad term for any shot where the camera follows a subject. If the camera follows the subject laterally, it is called a "trucking shot" instead (see next shot)
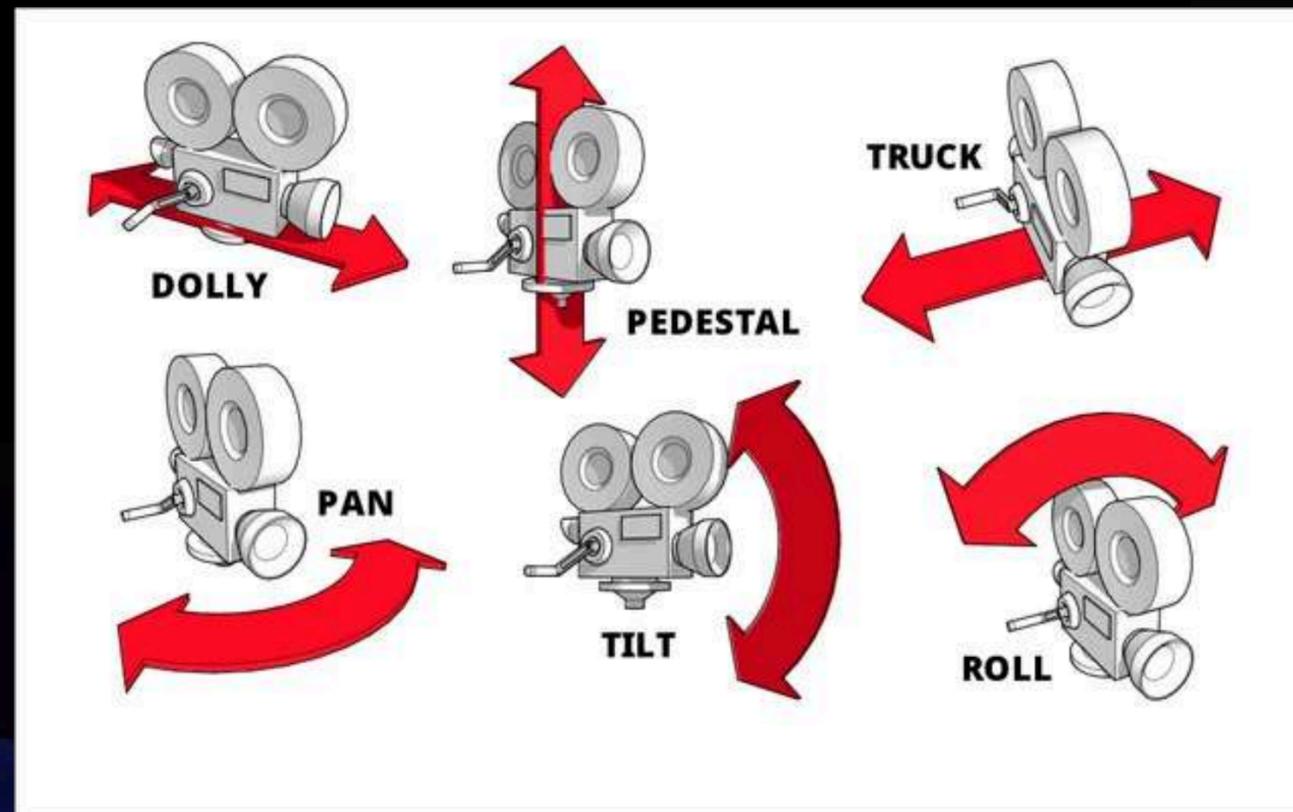

TRACKING

Camera Motions:

- **Truck:** the camera moves laterally (sideways) left or right. It's often used to walk alongside a character during a conversation.

THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Dolly Zoom (The "Vertigo" Effect): A brain-bending trick where you dolly the camera in while zooming the lens out (or vice-versa). The subject stays the same size, but the background appears to warp or shrink. It's used to show intense realization or vertigo. A slow Dolly In/Zoom out generates focus, intimacy, or rising tension. A Dolly Out/Zoom in instead generates a sense of abandonment, isolation, or revealing the "big picture."

DOLLY IN + ZOOM OUT

DOLLY OUT + ZOOM IN

THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Handheld Shot: a shot taken with the camera held in the operator's hands to keeps the natural shake, jitter, and instability of human movement and to convey a sense of intimacy or of adrenaline


HANDHELD

Camera Motions:

- **Static Shot / Fixed Shot:** a very common shot in which the camera does not move, tilt, pan, or zoom. It is mounted on a tripod or a stable surface and remains perfectly still throughout the entire duration of the take. With no camera movement to distract the eye, every tiny facial twitch or subtle gesture of the actor becomes magnified.

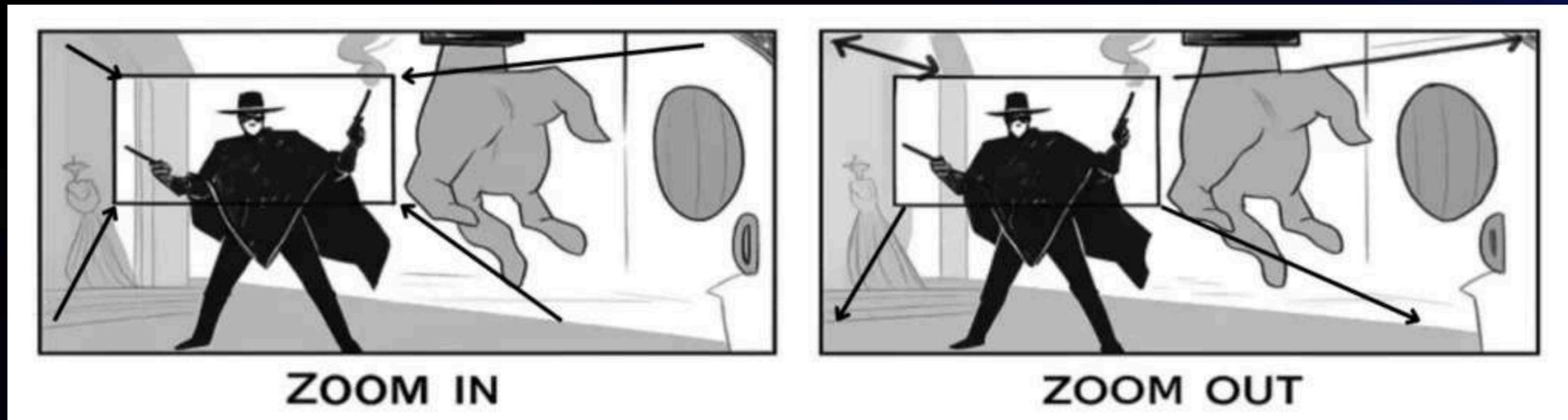# THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Arc Shot: the camera orbits around a subject in a circular or semi-circular path. It creates a sense of dizziness, heroics, or rising tension. Arcing around two people facing each other emphasizes the space between them and the "eye-to-eye" conflict. It's also a great way to reveal the world behind a character without cutting to a different shot.
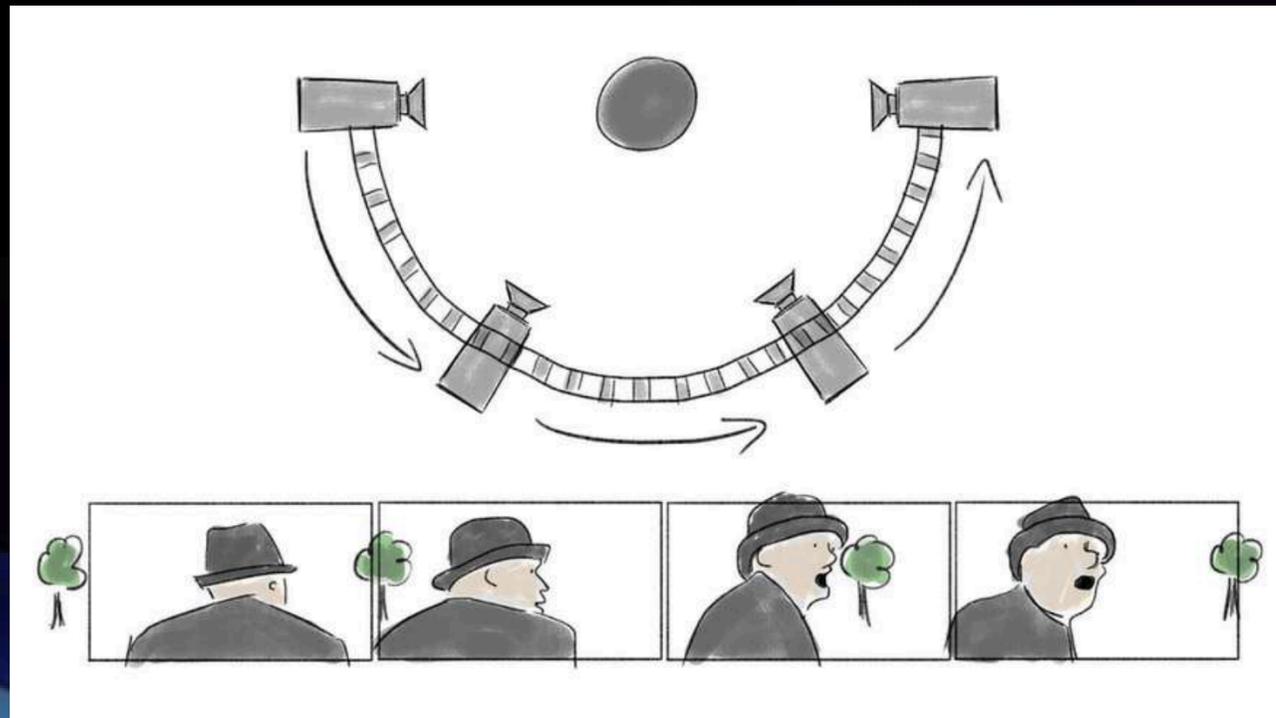
THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Zoom in/out: the focal length of the lens changes to make the subject appear closer or farther away. Unlike a dolly shot, the perspective of the background stays the same, which can sometimes feel "flatter". Only used in horror movies, or to convey a feeling of extreme humour or extreme drama. Do not use in our movie

THE GRAMMAR OF VISUAL STORYTELLING

Camera Motions:

- Slow Motion (slow-mo): the camera captures images at a much higher frame rate than they will be played back. A standard movie plays at 24 frames per second (fps); if the camera records instead at 120 fps, the action will appear five times slower when played back at the standard speed. In this way the audience can see every micro-expression on an actor's face during a moment of shock, sadness, or joy. It tells the viewer: "This moment is important; don't miss it" It is also ssed to make a character's entrance feel epic, or to see the trajectory of a bullet, the impact of a punch, or the grace of a jump. Slowing down the world makes it feel surreal, hazy, or romantic. It's also often used in flashbacks or drug-induced sequences.

Camera Motions:

- **Shot / Reverse Shot:** common motion for dialogues. One character is shown looking or speaking at another character (often off-screen), followed by a shot of the second character looking or speaking back at the first character.

THE GRAMMAR OF VISUAL STORYTELLING

Camera Focus:

- Deep Focus: The most common type of focus. In Deep Focus, every plane of the image, from the objects right in front of the lens to the distant mountains in the background, is in sharp focus. Best for: establishing shots, master shots, complex scenes where action is happening at different distances, and showing a character's relationship to a vast environment. It requires a good illumination

THE GRAMMAR OF VISUAL STORYTELLING

Camera Focus:

- Shallow Focus (or shallow depth of field) is a technique where only a small part of the image is sharp and clear (usually the foreground), while the background is softly blurred. It is one of the most powerful tools for directing the viewer's attention and it is very common for Close-ups shots

Camera Focus:

- **Tilt Shift** is an exteme version of Shalow Focus in which only a tiny portion of the screen is in focus, and the rest is blurry. The strange, localized blur feels like a memory or a dream where only one detail is clear.

TILT SHIFT

Camera Focus:

- **Rack Focus:** not a "motion" of the camera itself, but a shift in the lens's focus from one object in the foreground to another in the background (or vice versa). It's a classic way to show a character noticing something or to shift the power dynamic in a scene.

Light types:

All cinematography is a balance between light and shadows

- **Natural light:** Any light that occurs naturally: sunlight, moonlight, lightning strikes, fire light. It tends to be bluer than artificial light
- **Artificial light:** man-made light powered by electricity. It tends to be orange.

Light types:

- High Key Lighting: A low contrast ratio in which the background and the subject is lit evenly, with minimum contrast. Used for brighter, happier scenes

F THE RINGS: THE RETURN OF THE KING (2003)

JOJO RABBIT (2019)

Light types:

- **Low Key Lighting:** A high contrast ratio where the difference between light and darkness is higher, for maximum contrast. Used for darker, dramatic scenes

Light types:

- Hard Lighting: More directional light that creates stark shadows. It works best in more dramatic or intense scenes.

Hard Lighting

SE7EN (1995)

Light types:

- **Soft Lighting:** A more diffuse light that generates less defined shadows than hard lightnings. This type of lightning is often used for happier or romantic scenes. Some of the best naturally occurring soft lights comes during the "golden hour": just before sunset or after sunrise

Soft Lighting

THE SOCIAL NETWORK (2010)
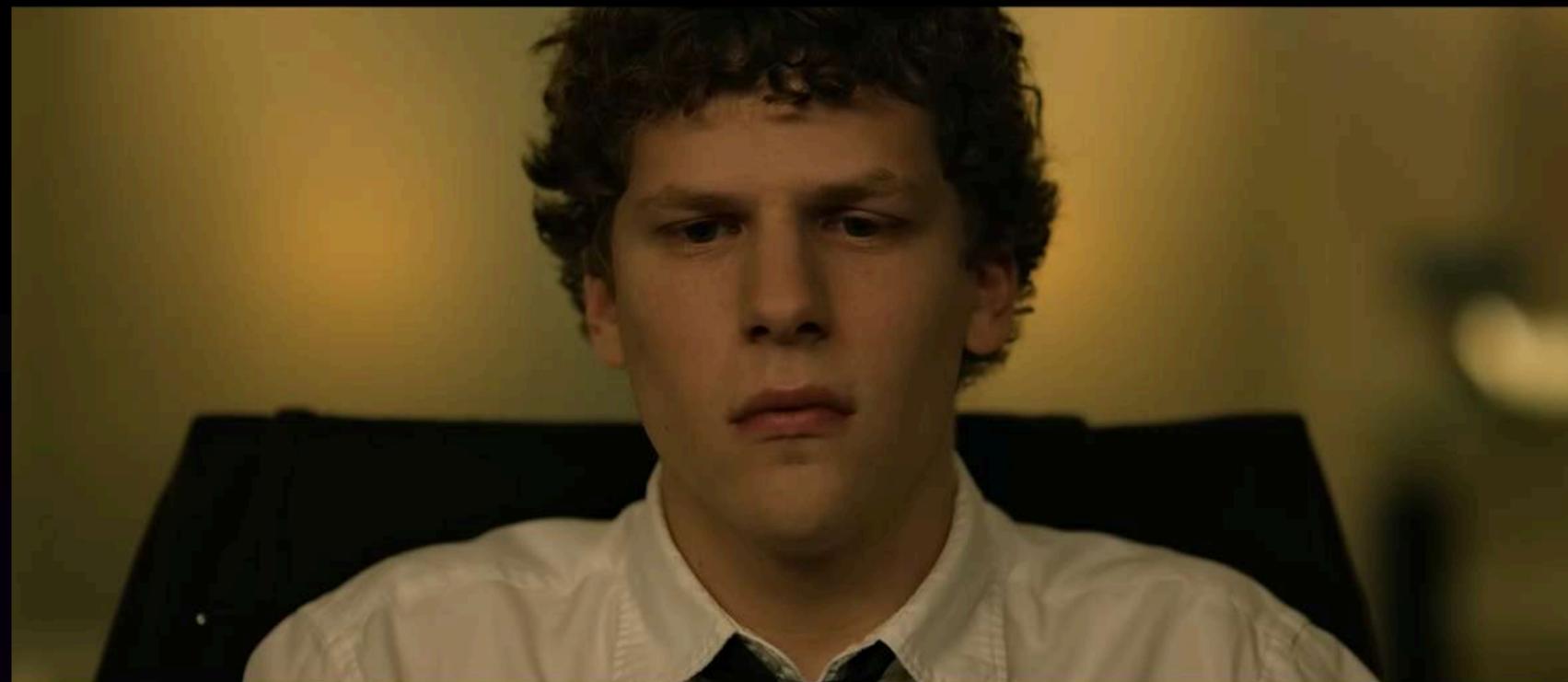
RESOURCES

Short guides on:

- Dolly Zoom: https://www.studiobinder.com/blog/best-dolly-zoom-vertigo-effect/
- Scene Transitions: https://www.youtube.com/watch?v=TKXBAaQB03U
- Shot Transitions: https://www.youtube.com/watch?v=FVR8zz8ci2k&t=30s
- Camera Shots: https://www.youtube.com/watch?v=AyML8xuKfoc
- Camera Framing: https://www.youtube.com/watch?v=qQNiqzuXjoM
- Camera Movements: https://www.studiobinder.com/blog/best-dolly-zoom-vertigo-effect/
- Camera Focus: https://www.youtube.com/watch?v=ZGNUqNFfD8M&t=5s
- Camera Rigs: https://www.youtube.com/watch?v=heJ9hWNb10g&t=296s
- Camera Lenses: https://www.youtube.com/watch?v=uSsIqR3DuK8
- Cinematic Lighting: https://www.youtube.com/watch?v=r2nD_knsNrc
- Cinematc Lighting part II: https://www.youtube.com/watch?v=N9mPQBZe06s
- Frame rates: https://www.youtube.com/watch?v=7lUFluLOh-s
- Composition: https://www.youtube.com/watch?v=hUmZldt0DTg&t=2s

OTHER RESOURCES

AI short movie winner of the 1-million dollar prize of the AI Film Award Festival:

- Lily: https://www.youtube.com/watch?v=f2d40eY7P1s&t=1s

Other AI movies that won the previous editions:
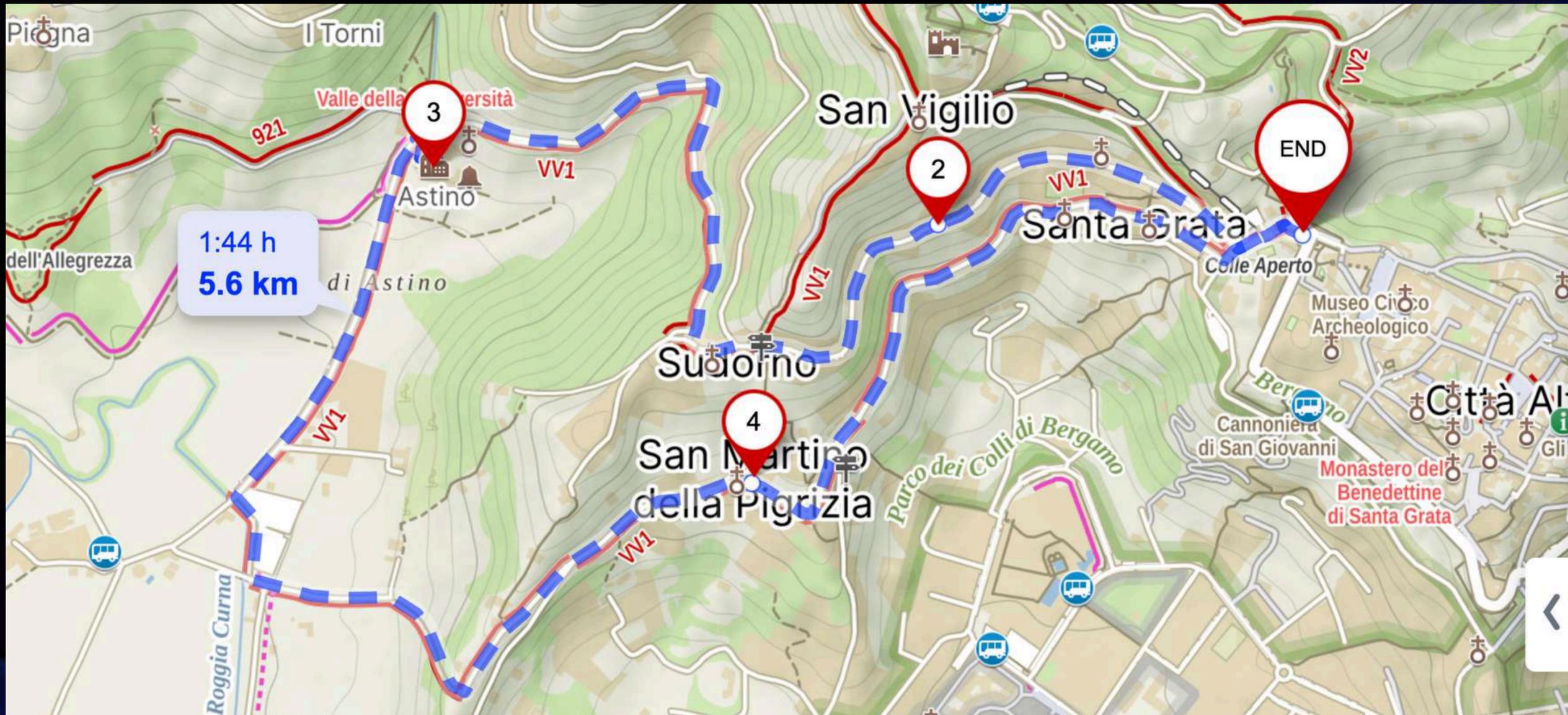
- https://www.filmawards.ai/

TOUR OF THE ASTINO VALLEY

The Astino valley is branded as the "most beautiful italian urban valley". If you live in Bergamo, you should do a tour of the valley along with the companions of your group

TOUR OF THE ASTINO VALLEY

You can start the walk from the bust stop of Colle Aperto, following this route:

TOUR OF THE ASTINO VALLEY

The Monastery of Astino in winter only opens during the weekends, so you probably won't be able to enter. Send me a photo of your group in front of the Monastery

TOUR OF THE ASTINO VALLEY

In case you find the Monastery open, take a look inside at the 1582 painting of the Last Supper of Alessandro Allori

PROJECT WORK #1

- 1000 members!
- Remove the lower part of your images before publishing them, if they have the logo of Gemini/ChatGPT/...
- Review the last facebook accounts in yellow to see if they are now recognized as human accounts
- Remember to publish 1 post per week and some comments
- Do not publish or comment during Saturdays or Sundays!

PROJECT WORK #2

- Group 4 and Group 5 now are the same group (#4)
- All students who have an english channel assigned just need to create the 300 long videos; the short videos can be uploaded from here: https://drive.google.com/drive/folders/1cGeRdfcvzibwzE9_wLWW09KEPmZNAqd8?usp=drive_link
- Check the new TikTok channels and their number of visualizations
- Add the first 300 long videos: videos longer than 8 seconds, made using HeyGen avatars. Each group'll generate one male face and one female face with the AI random face generator (https://perchance.org/ai-face-generator), and then convert the two faces to avatar videos using HeyGen (https://www.heygen.com/). Include audio, a background music of your choice and captions below each video (see the slides on avatar videos in the moodle of the course of IT). Long video numbers ends with the letter "a" to be able to list them together with short videos in numeric order
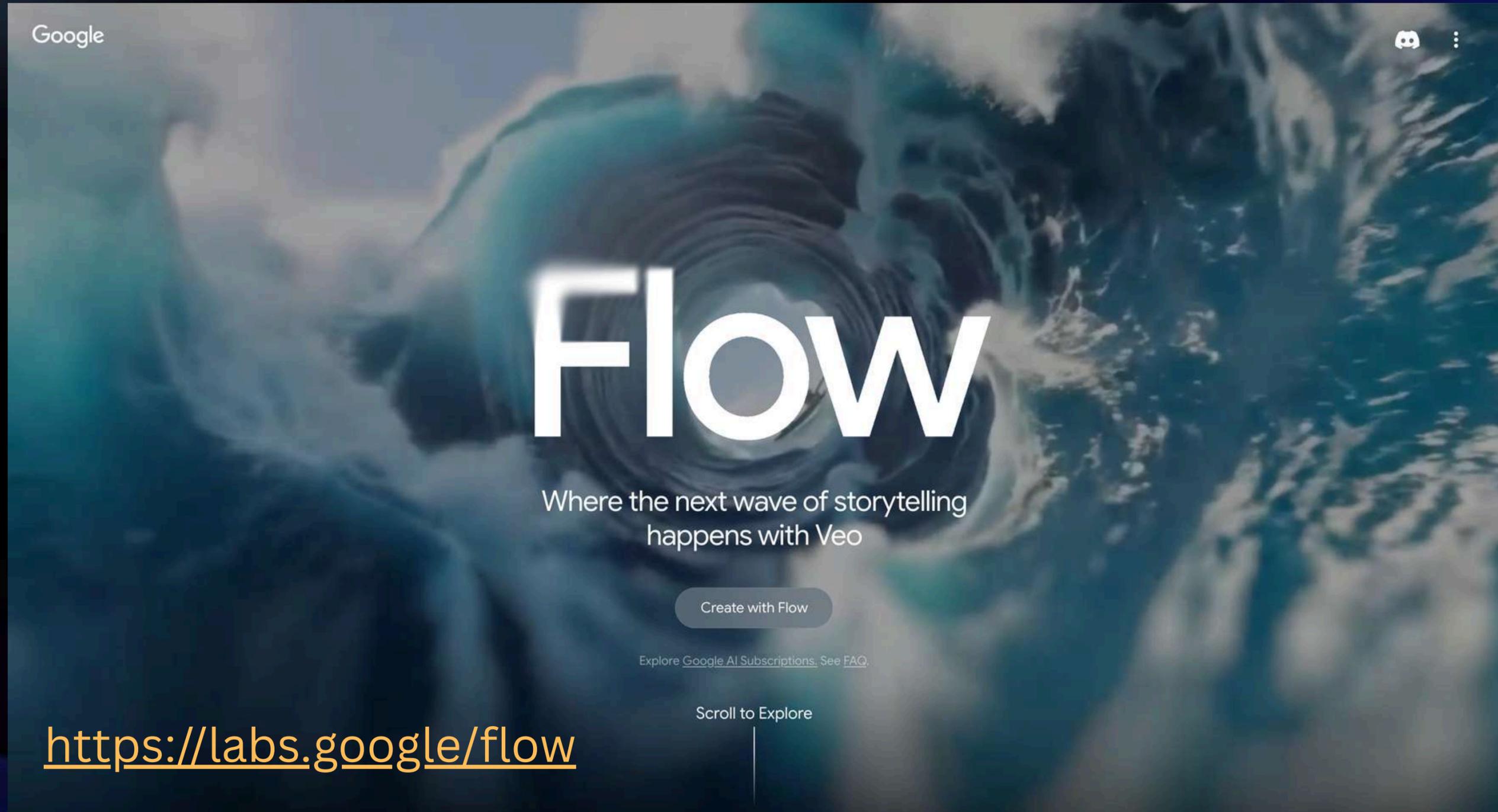
PROJECT WORK #3

- Listen to the new songs in the Moodle

PROJECT WORK #4

- Check if everyone was able to subscribe to Google AI Pro for students (access it with your email of the University)
- Check if everyone installed CapCut
- Read the Shabono Storyboard (in the moodle). Each group is in charge of generating certain Scenes, as shown in the Storyboard
- Show how to use Google Flow, in particular:
  - Extend and Jump to features of the Scenebuilder
  - Frames to Video and Ingredients to Video tools
  - Expand magic button
- Show how to use CapCut (example of Scene 10)

PROJECT WORK #4: GOOGLE FLOW

https://labs.google/flow

PROJECT WORK #4: GOOGLE FLOW

Main types of video prompts of Google Flow:

- Text to Video: use it only if you don't have names characters in the scene
- Frames to Video: use it to convert an image in a video, or to create a video shot between two existing video shots
- Ingredients to video: insert all the characters and items uploaded inside the video shot
- Extend (Scenebuilder): continue an existing shot, keeping the same characters and location of the previous shot
- Jump to (Scenebuilder): move to another location, but keeping the characters of the previous shot (however, it's only availabe to Veo 3.1 Quality that costs 100 credits per video, so it's better not to use it!)

PROJECT WORK #4: VEO 3.1 MAIN LIMITATIONS

Subscription type

- 1000 credits per month are not enough to generate all video clips we need, that's why we assigned each Scene to a different Group
- Cannot select the more advanced features to mantain consistency between faces, voices, dresses, musics and environment only available to Google AI Ultra users (features "Actor Gallery" and "Voice Seed")
- "Ingredients" tool of Google Flow is limited to uploading a max of 3 images
- Only Ultra users do not have the Veo watermark embedded the videos

Technical limitations

- Cannot generate a video clip longer or shorter than 8 seconds
- You can generate multiple video simultaneously (up to 4), but each video generation takes roughly one minute
- You cannot select the voices of the actors, so you must use a different tool (e.g: ElevenLabs)
- Video2video tool of Google Flow is quite limited yet: it can only remove or add objects, but not modify them

Intrinsic model limitations

- Veo 3.1 has issues at inserting exactly the number of objects you specified in the prompt
- Veo 3.1 doesn't always understand your prompt very well
- Veo 3.1 cannot generate videos of existing persons, children or naked bodies
- "Frames" tool of Google Flow sometimes generate a start frame that is slightly different from the last frame of the previous shot, and it is impossible to generate an intermediate frame between the two

PROJECT WORK #4: HOW TO CREATE A SCENE

- Read the original text of the Scene available in the Storyboard.
- Select the most important part of the Scene: the story that you want to show on the screen (you can't show everything written in the book!).
- Start selecting the dialogues. You can modify them, or shorten them if you think that this would improve the narration.
- Choose the best Google Flow video prompt type to make the shot
- Then, imagine the first shot. Who are the characters in the shot? What do they say? Where is the camera? At which angle? Does it move? Which is its focus? How is lightning?
- Introduce all the above description in the prompt, starting it with the sentence: "Cinematic, photorealistic, 35mm film grain."

PROJECT WORK #4: THINGS TO AVOID

- Never buy credits of Google AI Pro subscription
- Never generate videos with the models Veo 3.1 Quality or Veo 2 Quality, as they cost 100 credits each instead of 20
- Never generate a video with one or more characters using the Text to Video tool: upload instead their images in the Ingredients to Video tool, or use the Extend tool function of Google Flow's Scenebuilder to create consistent characters
- Never use the voices generated by Google Veo, as they are not consistent with previous shots: replace them in CapCut with the voices generated by ElevenLabs (the names of the voices to use of each characer are listed in column D of the Storyboard)
- Try to avoid too many static shots only with dialogues, as they become boring very quickly

TASKS FOR NEXT WEEK:

- Start writing the prompts of the first Shots of your first Scene, but without running them in Google Flow to save credits
- Decide with your group the two best male and female new avatars to use (plus their Elevenlabs voices). Then, start making a few long videos and upload them in your TikTok channel to see if they have more or less visualizations of the short videos
- If you have just created a new TikTok account, remember to modify your TikTok language preferences by setting the same language of your channel (TikTok main menu → Setting and privacy → Language → App language)

Suggested weekly AI movie: In Time (2011 dystopian movie). It shows you what may happen to a society in which people become immortal thanks to advanced technology

Next lesson we'll study Neural Networks and Deep Learning