

2.0 Objectives

In this session we'll be exploring in more depth what makes AI, how it's used, what it does, how you interact with it, some of the dangers, some of the benefits.

By the end of the module, you should

- Differentiate between automation and workflow
- Basic understanding of AI usage pitfalls
- Have completed a simple AI assisted content creation workflow

2.1 AI Overview November 2025

Should I use AI?

Yes. It's a tool that will not go away. A smart person will try and figure out if it's the right tool for the job. The wonderful podcast **The AI fix**  has an inspiring test - would you trust a well trained pigeon to do the job. If the answer is **yes**, then AI might be a good tool. If the answer is **Hell No!** then maybe you should trust your gut and find a better tool for your workflow.

What Sort of Workflows is AI good for?

Under the hood, all AI tools are constrained guessing machines. What?? Let's break that down:

- constrained ...
 - The neural networks predict what comes next based on training data
so ... If the likelihood of the training data being relevant to your problem is low - DO NOT USE AI
 - Neural networks work best when there is a lot of training data predicting a small number of outcomes in that data
so ... If your workflow squishes a large amount of data to a small number of outcomes, then AI might work e.g. *"Is this a picture of a dog?"*
but ... If your workflow takes sparse data and generates a lot of outcomes, **then AI will probably fail** e.g. *"write a story in Siwa Berber about tribal life"*
 - An LLM has billions of tokens and it tries to predict the next token based on the tokens it's already generated
so ... uniform, rich training data will likely give decent results e.g. *"write an iambic pentameter poem about broadcast workflow in English"*
but ... random, sparse training will likely give confident rubbish *"write an iambic pentameter poem about broadcast workflow in emoji"*
- guessing ...
 - this is *"just"* statistics and simplistic decision trees give bad outcomes. e.g. my training data for the new CEO role is based on successful CEOs in our sector (all of whom were men aged 50-70). How likely is it that the chosen candidate will be a globally successful woman aged 40?
 - best solution - fix the training data - very expensive and often the data does not exist
 - practical solution - place guard rails on the output to "have another go" if the output shows bias, errors or compliance problems
- constrained guessing
 - the practical solution to today's AI

AI Terms Glossary

Token

A fundamental unit of text that AI language models process, typically representing a word, part of a word, or punctuation mark.

Tokens are the building blocks that language models use to understand and generate text, with each token assigned a numerical identifier that the model can process. The tokenization process breaks down input text into these discrete units, which can vary in size from individual characters to complete words depending on the tokenization algorithm used. Understanding tokens is crucial because most AI models have token limits that constrain how much text they can process at once, and API costs are often calculated based on token usage. Different languages and writing systems tokenize differently, with some requiring more tokens than others for the same semantic content, which affects both processing efficiency and cost.

Model

A trained artificial intelligence system consisting of mathematical parameters and algorithms that processes inputs to generate predictions or outputs.

An AI model is the result of training neural networks on large datasets, encoding patterns and relationships discovered during training into billions of parameters that define how the system responds to new inputs. Models come in various architectures designed for different tasks, from transformer-based language models to convolutional networks for image processing, each optimized for specific types of data and problems. The size and complexity of a model, often measured in parameters, generally correlates with its capabilities, though architectural innovations and training techniques can dramatically affect performance. Once trained, models can be fine-tuned, quantized, or adapted for specific use cases, making them versatile tools that form the foundation of modern AI applications.

Context

The information and previous interactions that an AI model uses to understand and respond appropriately to the current input.

Context encompasses everything from the immediate conversation history to relevant background information, system instructions, and any retrieved documents that help the model generate accurate and coherent responses. The amount of context a model can handle is limited by its context window, measured in tokens, which determines how much information it can consider simultaneously when generating outputs. Managing context effectively is crucial for maintaining coherent multi-turn conversations, as models use context to understand references, maintain consistency, and build upon previous exchanges. Advanced techniques like context caching, summarization, and retrieval augmented generation help extend the effective context available to models beyond their native context window limitations.

Training

The process of teaching an AI model by exposing it to vast amounts of data so it learns patterns, relationships, and skills through iterative adjustment of its parameters.

Training involves feeding millions or billions of examples through a neural network while using mathematical optimization to adjust the model's parameters, minimizing the difference between predicted and actual outputs. This computationally intensive process requires enormous datasets, specialized hardware like GPUs or TPUs, and can take weeks or months for large models, consuming significant energy and computational resources. The training phase includes multiple stages, from initial pre-training on broad datasets to establish general capabilities, through fine-tuning on specific tasks or domains, to alignment training that shapes the model's behavior to be helpful and safe. The quality and diversity of training data profoundly impacts model capabilities, biases, and limitations, making data curation and preparation critical aspects of successful AI development.

Inference

The process of using a trained AI model to generate outputs or predictions based on new input data, distinct from the training phase.

Inference occurs when a model takes user input and applies its learned patterns to produce responses, generate text, classify images, or perform other tasks without modifying its underlying parameters. Unlike training which requires massive computational resources and time, inference is designed to be fast and efficient, often optimized through techniques like quantization, pruning, and specialized hardware acceleration. The cost and speed of inference directly impact user experience and operational expenses, with inference typically representing the majority of computational costs in deployed AI systems. Optimizing inference performance through model compression, batching requests, and efficient serving infrastructure is crucial for making AI applications practical and economically viable at scale.

Governance

The frameworks, policies, and practices that organizations establish to ensure responsible development, deployment, and use of AI systems.

AI governance encompasses ethical guidelines, risk management protocols, compliance requirements, and oversight mechanisms that guide how AI technologies are created and applied within organizations and society. This includes establishing accountability structures, defining acceptable use policies, implementing bias detection and mitigation strategies, and ensuring transparency in AI decision-making processes. Effective governance addresses concerns around data privacy, algorithmic fairness, safety, security, and the societal impacts of AI deployment, balancing innovation with responsible stewardship. As AI capabilities advance and become more integrated into critical systems, governance frameworks must evolve to address emerging challenges like autonomous decision-making, intellectual property considerations, and the need for human oversight in high-stakes applications.

Agentic AI

AI systems that can autonomously plan, make decisions, and take actions to achieve specific goals without constant human intervention.

Agentic AI represents a significant evolution beyond traditional AI systems by incorporating goal-directed behavior, autonomous decision-making, and the ability to interact with external tools and environments. These

systems can break down complex tasks into subtasks, create action plans, execute those plans using available resources, and adapt their strategies based on feedback. Unlike reactive AI that simply responds to inputs, agentic AI maintains context across interactions, learns from outcomes, and can even request additional information or resources when needed to accomplish objectives.

Large Reasoning Model (LRM)

Advanced AI models specifically designed to perform complex multi-step reasoning and problem-solving tasks that require logical thinking and strategic planning.

Large Reasoning Models extend beyond pattern recognition and text generation to focus on systematic thinking and logical deduction. These models are trained to decompose complex problems, maintain coherent chains of thought, verify intermediate steps, and arrive at well-reasoned conclusions. They excel at mathematical proofs, scientific reasoning, strategic planning, and debugging complex systems. Unlike standard language models that may hallucinate or make logical leaps, LRMs are optimized to show their work, validate assumptions, and acknowledge uncertainty when appropriate.

Vector Database

A specialized database system designed to store, index, and efficiently search high-dimensional vector embeddings that represent semantic meaning of text, images, or other data.

Vector databases solve the challenge of finding semantically similar information in large datasets by converting data into numerical vectors (embeddings) that capture meaning in high-dimensional space. When you search a vector database, it finds items whose vectors are closest to your query vector, enabling semantic search that understands context and meaning rather than just matching keywords. This technology is fundamental to modern AI applications, powering recommendation systems, similarity search, and enabling AI models to access relevant information from vast knowledge bases efficiently.

RAG (Retrieval Augmented Generation)

A technique that enhances AI language models by retrieving relevant information from external knowledge sources before generating responses, combining the benefits of search and generation.

RAG addresses the limitations of AI models being constrained to their training data by giving them access to up-to-date, domain-specific, or proprietary information. When a query is received, the system first searches a knowledge base (often using vector databases) to find relevant documents or passages, then provides these as context to the language model when generating a response. This approach reduces hallucinations, enables models to cite sources, allows for easy knowledge updates without retraining, and makes it possible to work with specialized information that wasn't in the model's training data.

MCP (Model Context Protocol)

A standardized communication protocol that enables AI models to securely access and interact with external data sources, tools, and services in a consistent way.

Model Context Protocol provides a unified framework for AI applications to connect with various external resources such as databases, APIs, file systems, and business tools without requiring custom integration code.

for each connection. It defines how models can discover available resources, request access to data, execute actions through tools, and maintain security boundaries. This standardization is crucial for building agentic AI systems that need to interact with enterprise systems, as it ensures consistent behavior, proper authentication, and controlled access to sensitive resources across different platforms and services.

MoE (Mixture of Experts)

A neural network architecture that uses multiple specialized sub-models (experts) with a routing mechanism that selectively activates only the most relevant experts for each input.

Mixture of Experts models achieve efficiency by dividing a large model into smaller, specialized components where each expert focuses on different types of patterns or domains. A gating network learns to route each input to the most appropriate experts, activating only a subset of the total parameters for any given task. This approach allows for much larger total model capacity while keeping computational costs manageable, since only a fraction of the model processes each input. The architecture is particularly valuable for handling diverse tasks or domains within a single model, as different experts can specialize in areas like mathematics, coding, creative writing, or specific languages.

ASI (Artificial Super Intelligence)

A hypothetical future form of artificial intelligence that would surpass human cognitive abilities across all domains, including creativity, problem-solving, and social intelligence.

Artificial Super Intelligence represents a theoretical milestone where AI systems would not just match but fundamentally exceed human-level intelligence in every meaningful way. Unlike narrow AI that excels at specific tasks or even general AI that matches human-level versatility, ASI would possess superior reasoning, learning speed, memory, creativity, and the ability to improve itself recursively. This concept remains firmly in the realm of speculation and future possibility, with significant debate among experts about whether ASI is achievable, what timeline might be realistic, and what safety measures would be necessary. The discussion around ASI drives important conversations about AI alignment, control mechanisms, and the long-term implications of increasingly capable AI systems.

LLM (Large Language Model)

Neural networks trained on vast amounts of text data to understand and generate human-like language, capable of tasks ranging from conversation to code generation.

Large Language Models represent a breakthrough in AI that emerged from scaling up transformer-based neural networks and training them on diverse text from the internet, books, and other sources. These models learn statistical patterns in language that enable them to generate coherent, contextually appropriate text, answer questions, summarize documents, translate languages, write code, and perform many other language-based tasks. The "large" refers to billions or even trillions of parameters that capture nuanced patterns in how language works. While LLMs can appear remarkably intelligent, they fundamentally operate by predicting likely text continuations based on patterns learned during training, rather than truly understanding meaning in the way humans do.

AI vs Machine Learning

Understanding the technologies shaping modern media

What is Artificial Intelligence?

AI is the broad field of creating machines that can mimic human intelligence.

It encompasses reasoning, learning, problem-solving, perception, and language understanding.

Key Characteristics

Scope: Broad concept of intelligent machines that can perform tasks requiring human-like intelligence

Goal: Simulate human intelligence to solve complex problems

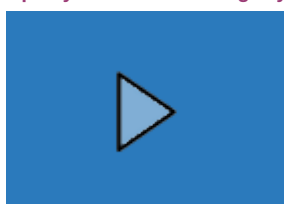
In Media: Script analysis, audience prediction, visual effects, casting decisions

Example: IBM Watson analysing scripts for Sony Pictures

Key Concept: AI is the umbrella term that covers all intelligent systems, including machine learning, robotics, expert systems, and natural language processing.

Learn More:

- **YouTube:** Two Minute Papers - simple how it works videos
 - from 2015 - Artificial Neural Networks and Deep Learning - <https://youtu.be/rCWTOOgVXyE>



- **YouTube:** 3Blue1Brown - deeper dives for deeper learning
 - from 2015 - But what is a neural network? | Deep learning chapter 1 - <https://youtu.be/aircAruvnKk>



- **Guide:** Google Cloud AI vs ML - cloud.google.com/learn/artificial-intelligence-vs-machine-learning 

What is Machine Learning?

ML is a subset of AI where systems learn from data without explicit programming.

Algorithms improve performance through experience and pattern recognition in data.

Types of Machine Learning

1. **Supervised Learning:** Algorithm learns from labelled training data
 - e.g., predicting box office success based on historical data
2. **Unsupervised Learning:** Algorithm finds patterns in unlabelled data
 - e.g. audience segmentation
3. **Reinforcement Learning:** Algorithm learns through trial and error with rewards
 - e.g. optimizing video editing workflows

Key Characteristics

Scope: Subset of AI focused specifically on learning from data

Goal: Learn patterns from data to make predictions or decisions

In Films: De-aging actors, automated editing, VFX creation, audience analytics

Example: Netflix recommendation algorithms, ILM's de-aging technology

Learn More:

- **YouTube:** StatQuest with Josh Starmer - <https://youtube.com/@statquest> 
- **YouTube:** freeCodeCamp Machine Learning courses - youtube.com/@freecodecamp 
- **YouTube:** Sentdex channel - youtube.com/@sentdex 
- **Guide:** Google Machine Learning Crash Course - developers.google.com/machine-learning/crash-course 

Key Differences: AI vs ML

Aspect	Artificial Intelligence	Machine Learning
Scope	Broad concept of intelligent machines	Subset of AI focused on learning from data
Goal	Simulate human intelligence	Learn from data patterns
Approach	Uses various techniques including ML, rules, logic	Uses algorithms to find patterns in data
Film Example	IBM Watson analyzing entire scripts	Netflix learning viewing preferences
Relationship	The umbrella category	A tool within AI - specific learning for a specific purpose

 **Simple Analogy:** If AI is the entire film industry, then Machine Learning is the rigging department within visual effects—a specialized and powerful subset that handles specific types of intelligent tasks.

AI & ML in Media

Industry Impact

70% of movies now integrate AI technology during production

Source: market.us AI in Film Market Report (2025)

Real-World Applications in Scripted Content

De-aging Technology

Machine learning algorithms enable realistic age transformation of actors on screen.

Examples: Luke Skywalker in *The Mandalorian*, Will Smith in *Gemini Man*, Tom Hanks in *Here*

Script Analysis

AI systems predict box office success and audience reception with up to 90% accuracy.

Tools: ScriptBook (Sony Pictures), Cinelytic (Warner Bros.), 20th Century Fox Merlin system

Visual Effects & CGI

ML powers automated rotoscoping, digital character creation, and realistic environments.

Examples: Thanos in *Avengers*, digital environments, automated background removal, facial replacement

Audience Prediction

Algorithms analyze viewer preferences to optimize content recommendations and marketing.

Platforms: Netflix recommendation engine, streaming platform analytics, targeted advertising

Additional Applications

- **Post-Production:** Automated video editing, color grading, sound mixing
- **Pre-Production:** Casting decisions, location scouting, budget optimization
- **Production:** Real-time VFX previews, intelligent camera systems, scene composition
- **Marketing:** Trailer generation (IBM Watson for *Morgan*), targeted promotions, release strategy
- **Content Creation:** AI-assisted scriptwriting, synthetic voiceovers, music composition

 **Industry Insight:** AI is not replacing human creativity—it's augmenting it. Filmmakers use these tools to handle repetitive tasks, explore new creative possibilities, and make data-driven decisions while maintaining artistic control.

Explore More:

- **YouTube:** Filmmakers Academy - [AI Video Tools for Filmmakers](#) 
- **Article:** SmartClick AI - [How AI is Used in the Film Industry](#) 
- **Blog:** Ultralytics - [AI in Filmmaking: New Creative Possibilities](#) 
- **Research:** market.us - AI in Film Market Report (2025) - Industry valued at \$14.1 billion by 2033
- **Article:** Fortune Magazine - [How AI is Reshaping the Way Movies are Made](#) 

Key Takeaways

- AI is the broad field; ML is a powerful subset focused on learning from data
- Both technologies are already integral to modern filmmaking workflows
- These tools augment human creativity rather than replace it
- Understanding AI/ML helps filmmakers leverage these technologies effectively
- The film industry continues to innovate with AI applications across all production stages

This guide provides an introduction to AI and Machine Learning for filmmakers. For deeper technical knowledge, explore the cited resources and continue learning through hands-on experimentation with AI tools in your own projects.

2.4 Free Tools Directory

Free AI Tools Directory (2025)

This comprehensive guide lists free AI tools across six key categories. All tools listed offer free tiers or trials, making them accessible for individual developers, creators, and small teams to explore AI capabilities without financial commitment.

1. Chatbots for Conversational Learning & Work

These AI assistants help with learning, problem-solving, and productivity through natural conversation.

ChatGPT (Free Tier) - chatgpt.com 

OpenAI's conversational AI with access to GPT-3.5. Handles writing, analysis, coding help, and general knowledge questions. Free tier includes unlimited messages with some rate limiting.

Google Gemini - gemini.google.com 

Google's AI assistant integrated with Google services. Provides real-time web search capabilities, supports multiple languages, and connects with Docs and Gmail.

Microsoft Copilot - copilot.microsoft.com 

Built into Edge browser and Microsoft Office. Uses OpenAI technology with plugins and real-time browsing. No signup required for basic access.

Claude (Free Tier) - claude.ai 

Anthropic's AI assistant focused on helpful, harmless, and honest responses. Strong at analysis, coding, and creative writing. Free tier available via web interface.

QuillBot AI Chat - quillbot.com/ai-chat 

Conversational AI specialized in writing tasks. Free tier allows 20 queries per day. Good for content generation, summarization, and text improvement.

2. Agents That Can Search the Internet for Results

These AI agents actively search and retrieve current information from the web to answer queries.

Perplexity AI (Free Tier) - perplexity.ai 

AI-powered answer engine that searches the web and provides cited, real-time responses. Free tier includes basic searches with source citations. Excels at research and fact-finding.

Perplexity Comet Browser (Free) - perplexity.ai/comet 

AI browser with sidecar assistant that helps answer questions about web pages, summarize content, and navigate sites. Now available free worldwide.

Perplexity Deep Research (Free Limited) - perplexity.ai ➞

Conducts in-depth research by performing dozens of searches and reading hundreds of sources autonomously. Free users get limited queries per day.

Microsoft Copilot - copilot.microsoft.com ➞

Includes real-time web search integration, providing current information and links to sources directly in responses.

Success Criteria: A good internet-searching AI agent should: (1) Cite sources clearly, (2) Access current information (not just training data), (3) Distinguish between factual results and synthesis, (4) Provide links for verification, and (5) Handle ambiguous queries by refining searches.

3. Agents That Can Write Code

AI coding assistants that generate, complete, and debug code across multiple programming languages.

GitHub Copilot (Free Tier) - github.com/features/copilot ➞

AI pair programmer from GitHub/OpenAI. Free tier includes code completion, chat assistance, and limited access. Integrates with VS Code, Visual Studio, and JetBrains IDEs. Over 15 million users.

Cursor (Free Plan) - cursor.com ➞

AI-native code editor built on VS Code. Free plan includes limited requests. Features Agent mode for autonomous coding, multi-file editing, and codebase context understanding.

Windsurf - codeium.com/windsurf ➞

AI coding assistant with innovative "Cascade" collaboration features. Offers smooth AI-human teamwork with real-time context awareness. Free tier available.

Replit AI - replit.com ➞

Cloud-based IDE with built-in AI assistance. Helps with code generation, debugging, and learning. Free tier includes basic AI features.

Amazon CodeWhisperer - aws.amazon.com/codewhisperer ➞

AI code generator with real-time suggestions. Includes security scanning and supports multiple languages. Free tier available for individual developers.

Microsoft IntelliCode - visualstudio.microsoft.com/services/intellicode ➞

Free AI-assisted development for Visual Studio and VS Code. Provides context-aware code completions trained on open-source projects.

4. Generative AI for Images

Text-to-image generators that create visuals from written descriptions.

DALL-E 3 (via ChatGPT Free) - chatgpt.com ➞

OpenAI's latest image generation model integrated into ChatGPT. Free tier allows 3 images per day. Strong prompt understanding and natural language processing.

Bing Image Creator (Free) - bing.com/images/create ➞

Microsoft's implementation of DALL-E technology. Free to use with daily generation limits. Integrated with Bing Chat and available on web, Android, and iOS.

Stable Diffusion (Open Source) - stability.ai ➞

Free, open-source image generation model. Can run locally with 8GB+ VRAM or use via web platforms. Highly customizable with community-created extensions.

Leonardo.AI (Free Tier) - leonardo.ai ➞

Specialized in high-quality game assets and character portraits. Free tier includes limited daily generations. Good for detailed character and environment creation.

Adobe Firefly (Free Tier) - firefly.adobe.com ➞

Adobe's AI image generator with commercial-safe training. Free tier includes monthly generation credits. Integrates well with Creative Cloud apps.

Canva Text-to-Image (Free Tier) - canva.com ➞

Simple AI image generator built into Canva. Free tier includes basic generations. Perfect for quick blog graphics, social media, and marketing materials.

Dream by Wombo (Free) - dream.ai ➞

Mobile-first AI art generator with unique artistic styles. Free with ads. Popular for creative and abstract imagery.

5. Generative AI for Video

AI tools that generate video content from text prompts, images, or existing video clips.

Runway (125 Free Credits) - runwayml.com ➞

Industry-leading AI video generator. Free tier includes 125 one-time credits. Features Gen-4 model for high-quality generation, plus 30+ AI magic tools for video editing.

Pika.art (Free Tier) - pika.art ➞

Latest 2.2 model offers 10-second 1080p videos. Free tier available. User-friendly with strong anime-style generation and community features.

PixVerse (Free Tier) - pixverse.ai ➞

Social media sensation with v4.5 model. Free tier with 5-second video generation. Popular for viral effects like AI Kiss, AI Hug, and AI Fighting templates.

Haiper (100 Free Credits) - haiper.ai ➞

\$10/month gets 1,500 credits, but 100 credits free on signup. Fast generation with text-to-video, image-to-video, and creative templates. Mobile app available.

Luma Labs Dream Machine (Free Tier) - lumalabs.ai/dream-machine ➞

Beautiful interface with Ray3 model in 4K. Free tier available. Known for aesthetic quality and smooth camera movements, though physics can struggle with fast action.

Google Veo 3.1 (Free Trial via Gemini) - gemini.google.com ➞

Google's video generation model. Limited free access through Google Gemini. High-quality realistic video generation with good prompt following.

6. Generative AI for Audio

AI tools for creating music, sound effects, and voice content.

Music Generation

Suno (Free Tier) - suno.com ➞

Leading AI music generator. Free tier with v4.5 All model creates full songs with vocals and lyrics from text prompts. Supports multiple genres and custom lyrics.

Udio (Free Tier) - udio.com ➞

Collaborative AI music studio. Free tier available. Strong vocal quality and realistic song generation. Good for teams and creators wanting full tracks.

Boomy (Free Tier) - boomy.com ➞

Beginner-friendly instant song creation. Free tier allows music creation and distribution to Spotify and Apple Music with potential royalties.

Mubert (Free Tier) - mubert.com ➞

Generative music for background use. Free tier perfect for streamers, livestreams, and content creators. Offers API access for developers.

AIVA (Free Tier) - aiva.ai ➞

AI composer specializing in classical and cinematic music. Free for personal use. Creates emotional soundtracks and orchestral pieces.

Sound Effects Generation

ElevenLabs Sound Effects (Free Tier) - elevenlabs.io/sound-effects ➞

Text-to-sound-effect generator. Free tier includes multiple daily generations. Creates royalty-free SFX for videos, games, and media projects.

FineVoice AI Sound Effect Generator (Free Tier) - finevoice.fineshare.com/ai-sound-effect-generator ➞

Video and text to sound effects. Free tier available. Synchronized audio technology for video projects with copyright-free outputs.

SFX Engine (Free Credits) - sfxengine.com ➞

Unlimited unique sound effects with detailed text descriptions. Pay-as-you-go model with initial free credits. Commercial use license included.

Adobe Firefly Sound Effects (Free Tier) - adobe.com/products/firefly/features/sound-effect-generator ➞

Generate sound effects from text, voice recordings, or act them out into your microphone. Free tier with monthly credits. Integrates with Adobe workflow.

Canva Sound Effects Generator (Free Tier) - canva.com/features/ai-sound-effect-generator ➞

Simple text-to-sound-effect tool built into Canva. Free tier includes basic generations. Good for quick social media and video projects.

7. HuggingFace: The Community AI Toolkit

HuggingFace  is a comprehensive platform that hosts over 500,000 free AI models, datasets, and interactive applications called "Spaces." Unlike the commercial tools above, HuggingFace is a community-driven platform where researchers and developers share their AI models and applications. It's completely free for basic use and provides access to cutting-edge AI tools that you can use directly in your browser or integrate into your own projects.

What Makes HuggingFace Special?

Models: Pre-trained AI models for virtually any task (text, image, video, audio)

Datasets: Thousands of datasets for training and testing

Spaces: Interactive web apps you can use immediately without coding

Free to Use: Most tools are completely free with no signup required

Open Source: Code is available for learning and customization

How to Find Specific Tools on HuggingFace

HuggingFace Spaces makes it easy to find AI tools for specific tasks. Here's how to navigate the platform:

Step-by-Step Guide:

- 1. Visit HuggingFace Spaces:** Go to huggingface.co/spaces 
- 2. Browse by Category:** Click on categories like "Video Generation," "Text Analysis," "Image Captioning," or "Speech Synthesis" to filter results
- 3. Use the Search Bar:** Type keywords describing your task (e.g., "video subtitle," "image caption," "voice clone")
- 4. Check Popularity:** Look for tools with high engagement (hearts ♥, running status, and community usage)
- 5. Try It Instantly:** Most Spaces let you upload your file and get results immediately in your browser

Example 1: Adding Captions/Subtitles in Native Language to Video

Whisper Auto-Subtitled Video Generator - **HuggingFace Space**

What it does: Input a YouTube video link and automatically generate subtitles in the video's original language or translate to English.

How to use: Paste a YouTube URL, select language options, and the tool processes the video and generates downloadable subtitled video.

Technology: Uses OpenAI's Whisper model for speech recognition and transcription.

Best for: Content creators, educators, and accessibility needs

Generate Subtitles for Videos - **HuggingFace Space**

What it does: Upload video or audio files (or provide URLs) and generate subtitles in SRT format for multiple languages.

How to use: Upload your file, select target language, and download the generated SRT subtitle file.

Output format: Standard SRT subtitle files compatible with all video players and editors

Best for: Multilingual content, international audiences, video editing workflows

Example 2: Video Captioning (Describing Video Content)

NVIDIA Describe Anything - [HuggingFace Space](#)

What it does: Generates detailed, localized descriptions of specific regions in images and videos.

How to use: Upload an image or video, click on regions of interest, and receive detailed AI-generated descriptions.

Use cases: Video accessibility, content indexing, automated video metadata, scene understanding

Best for: Media professionals, researchers, accessibility applications

Video Captioning Tool - [HuggingFace Space](#)

What it does: Automatically generates descriptive captions for video content using AI vision models.

How to use: Upload a video file and receive AI-generated textual descriptions of the visual content.

Best for: Video cataloging, search optimization, content management systems

Tips for Using HuggingFace Successfully

Start Simple: Try a few Spaces first before diving into models and code

Read the Description: Each Space has instructions and limitations

Check Upload Limits: Free tiers may have file size or processing time limits

Explore Alternatives: If one tool doesn't work well, search for similar ones - there are often multiple options

No Account Needed: Most Spaces work without creating an account, though some features require login

Community Support: Use the discussion tabs to ask questions or report issues

Other Useful HuggingFace Categories

- **Text-to-Speech:** Convert text to natural-sounding voices in multiple languages
- **Voice Cloning:** Create custom voice models from audio samples
- **Object Detection:** Identify and locate objects in images and videos
- **Background Removal:** Automatically remove backgrounds from images
- **Image Upscaling:** Enhance image resolution with AI
- **OCR (Text Extraction):** Extract text from images and documents
- **Language Translation:** Translate between dozens of languages
- **Sentiment Analysis:** Analyze emotional tone in text
- **Code Generation:** AI assistants for programming tasks

Start Exploring - [Browse All HuggingFace Spaces](#)

Visit the Spaces homepage to discover hundreds of thousands of free AI tools. Use filters to find exactly what you need, or browse trending Spaces to see what the community is building.

Important Notes

Free Tier Limitations: Most tools have usage limits on free tiers (daily generations, credits, or rate limits). Check individual platforms for current limitations.

Commercial Use: Some free tiers restrict commercial use. Review terms of service if planning commercial projects.

Rapid Evolution: The AI tool landscape changes quickly. New features and pricing models are frequently updated throughout 2025.

Quality Variance: Free tiers often use older or less powerful models compared to paid subscriptions, though many still produce excellent results.

Published

Assign To

Edit



Try some free tools from huggingface

Student Pairs - assignment a2.6

1, 13
14, 12
15, 11
16, 10
2, 9
3, 8
4, 7
5, 6

Choose a new partner and set your own task to explore a free tool on huggingface. You can ask ChatGPT for help on where to start. Same example prompts:

- How would I use huggingface to put tools together to turn an uploaded photo into a 10 second video with audio
- How would I use huggingface to scan a picture with cars in it and list the types of cars it finds
- How would I use huggingface to add captions to a video

Refer back to the [Free Tools page](#) and make something awesome with your partner

Important Notes

Free Tier Limitations: Most tools have usage limits on free tiers (daily generations, credits, or rate limits). Check individual platforms for current limitations.

Commercial Use: Some free tiers restrict commercial use. Review terms of service if planning commercial projects.

Rapid Evolution: The AI tool landscape changes quickly. New features and pricing models are frequently updated throughout 2025.

Quality Variance: Free tiers often use older or less powerful models compared to paid subscriptions, though many still produce excellent results.

Points 0

Submitting a text entry box or a website url

Due	For	Available from	Until
-	Everyone	-	-

 [Rubric](#)

2.7 Bad AI - the Siwi experiment.

I asked one AI to write a story and another AI to critique it. I asked them both to think hard about it.

prompt Write a short modern story about a young person engaging with tribal elders in Siwi, with an English translation. The Siwi should be understandable by Berber natives.

Story in Siwi (Tasiwit)

Amzwaru n *Yasmina* i s-gʷra d urar n Isiwan. Yasmina d ɣur-t tant n tyʷart, yesεaʒ ga-mtawet s tazrawt n imdyazen-is. Imɣarnen n tmurt-is yettεazen d imɣanen n wass n tmurt, gar tamaɣna n ɣurba azzay.

Ur s-yufa miɣ, Yasmina y-ixdem ɣef tasunt agg ayt n wazurgaren. Yewwi ɣef umennu-n tiɣran: "Am ad awal iw iwata ma ykra wi?" Yeshen wazurgar, am ɣhayef: "Yasmina, awal nk iw am yemsaw, imma ttik is yim n wenza. Tzemred akk-ur gʷrani ad -gh ɣer tazrawt-nni. Imeskar n yiman-nni yella d amjaɣ."

Yasmina y-smeed iɣtiram s wazurgar, is-tfaɣ s wayn m-yessawal. Tazrawt-nni tella tazriɣa iwul, u yella yimsawen n wawal, agi d tislit n tmurt-is, d tmaziɣt n ixefraren. Yasmina tesεed i uzurgar ad yeddu tadist n tazzumt: ad yefka yined n yiman-nsen, ad yeddu aɣrim n tlawin-is, ad yefka iɣer n tusna n tesga.

Mi y-sufar ɣer tammarut n tmurt agg, wazurgar y-iwazzi Yasmina: "Yasmina, k-tta zɣur iwaten. Ma tsn kw iw am tanzdirt n tmazirt-nni. Ad yettɣfen awal-nwen s yiman-nwen, ad yeseed wen d-amjaɣ."

Yasmina tesεed, tixsi issin iwulas: "Am urgul ad adwel gʷra d urar n Isiwan, d-tafukt n tmurt-nwen, d-izri n tleɣrart n ixefraren."

Wazurgar yerna d-asirem: "Nekk d ayaɣ, Yasmina. Tazzumt n tmurt-nni d ayla, d-amjaɣ ayen ghaf yiman-nwen."

Yasmina tkemmel ayt n tezdag igez, tezmer d amɣakɣ: minden n tmurt-nni yella d tawiza n ɣef imdyazen-nwen — tazrawt agg d tazmit n uzaɣar, d-taɣallit n tmazirt-nni ur yettakkaɣ d ifassen.

English Translation

Young **Yasmina** lives in the heart of the Siwa Oasis. Yasmina is the daughter of a modest family, and she carries the hope of her ancestors in her proud eyes. The people of her land speak with deep respect and the winds of the desert carry the stories of their history.

Every morning, Yasmina walks to meet the elders of her tribe. She approaches the gathering place of the old ones and asks gently, "Elder, may I speak with you about my future?" The elder, calm and wise, replies: "Yasmina, your words are precious, like water in the desert. You must stay grounded and listen deeply. The strength of your heart lies in your people. The wisdom of your mind is your heritage."

Yasmina listens respectfully to the elder, and she opens her heart when she speaks. The gathering place becomes alive, and many voices join in her words — it is a conversation about their land, about their Berber identity, their responsibility to protect their heritage. Yasmina helps the elders to weave a plan: to pass down their songs, to guard the stories, to plant new palms beside the old ones.

When the sun begins to dip low in the sky, the elder smiles at Yasmina: "Yasmina, you are the one who returns home. You carry the memory of this land in you. Speak your truth with your mind, help your tribe with your heart."

Yasmina nods, her eyes shining: "I promise to live in the heart of Siwa, as a light for this land, as a guardian of its heritage."

The elder says with hope: "Yes, Yasmina, this land's future lies in your hands, in the hearts of our children."

Yasmina finishes her speech with determination, carrying the pride: the memory of this land is now a bridge between generations — this gathering place is a promise of renewal, a bond of Tamazight and the desert that will not be broken.

prompt Check the Siwi text for more authenticity (vocabulary, grammar) .

Some Linguistic Observations & Suggestions

1. Name of the language

- Actually, native Siwi speakers generally call their language **Siwi**, or **Jlan n Isiwan** ("speech of Siwa").
lughat.blogspot.com+1 ➞

- Using “Tasiwit” is less appropriate for Siwi, because the “t-...-t” naming pattern isn’t typical in Siwi. lughat.blogspot.com+1 ↗

- So, in your story, it’s more “Siwi” than “Tasiwit”.

2. Grammar / Word Order

- According to sources, Siwi tends to favor **Subject-Verb-Object** (SVO), not necessarily VSO. en.wikipedia.org ↗
- Your sentences sometimes feel more like a Berber from the Maghreb (or a generic Tamazight) style. Some phrasing might need to be simplified or re-ordered to feel more Siwi-native.

3. Lexicon / Borrowings

- Siwi has **significant Arabic influence**, especially from Egyptian and Bedouin Arabic. endangeredvoices.org+1 ↗
- Some of the words in the story (like “imdyazen,” “tazmit,” “tazrawt”) might not reflect typical Siwi vocabulary or morphological patterns. Without very detailed Siwi lexicons, it’s hard to be exact, but one should favor simpler Berber roots + perhaps Arabic borrowings for modern concepts.

4. Morphology of Nouns / Gender / Affixes

- Siwi nouns typically have Berber gender (masculine/feminine) and number markers. en.wikipedia.org ↗
- For example, feminine **-t** is common in Berber. Using “tazrawt” (for “gathering / place”) is okay in principle, but we should make sure consistency in gendered forms is maintained.

5. Pronouns / Possessives

- According to Siwi grammar sources, pronouns and suffixes are quite specific. de.wikipedia.org+1 ↗
- Make sure that possessive constructions (“my future”, “our land”) use suffixes that match Siwi patterns.

6. Cultural Authenticity

- It’s good to use words that evoke **Siwa Oasis life**: maybe mention *dates (titini)*, *palm trees*, *salt lakes*, *oasis winds*. Using some botanical vocabulary can reinforce “Siwi-ness.” (There is a Siwi botanical lexicon for cultivated plants. [Vincent Battesti](#) ↗)
- Also, “tribal elders” is fine, but in Siwi context more likely “wise men / old ones of the oasis” / “wazuṛgar” could be acceptable, though I’m not certain “wazuṛgar” is attested exactly in Siwi — it’s a Berber root but might differ in Siwi.

2.8 Worse AI - is it a bubble?

There are some worrying data points about AI:

1. 2025-02-21 [Tech Crunch](#) ⇨ **Court filings show Meta staffers discussed using copyrighted content for AI training**

- "My opinion would be (in the line of 'ask forgiveness, not for permission'): we try to acquire the books and escalate it to execs so they make the call," wrote Xavier Martinet, a Meta research engineer
- [Libgen](#) ⇨ is a "links aggregator" that provides free access to copyrighted works from publishers
- "Libgen essential to meet State Of The Art numbers across all categories," - Sony Theakanath, director of product management at Meta,

2. 2025-09-26 [Irish Times](#) ⇨ **Top companies keep talking about AI – but can't explain the upsides**

- Financial Times Analysis of hundreds of filings suggest the S&P 500 businesses are clearer about the risks than benefits.
- Microsoft, Alphabet, Amazon and Meta extol AI's benefits and are pledging to invest \$300 billion in 2025 for infrastructure.
- Most anticipated benefits, such as increased productivity, were vaguely stated and harder to categorise than the risks.
- The most commonly cited concern was cybersecurity, which was mentioned as a risk by more than half of the S&P 500 in 2024.
- MIT Media Lab found 95 per cent of generative AI pilots in the workplace failed

3. 2025-10-07 [CyberSecurityDive](#) ⇨ **Public disclosures of AI risk surge among S&P 500 companies**

- More than 70% companies on the S&P 500 - "AI is a material risk" - ⇨ [The Conference Board](#) ⇨.
- 38% worry about reputational risk
- 20% about cybersecurity

4. 2025-10-12 [Medium](#) ⇨ **AI Just Had Its "Big Short" Moment**

- No AI company or AI infrastructure is profitable. They are all losing money hand over fist.
- They are using equity financing (selling shares of the company) or debt financing (borrowing money) to keep the lights on.
- Each generation of model needs much-much more compute for a marginal improvement in performance
- Deutsche Bank actually calculated the extent of the AI bubble & found half the current USA growth can be attributed to AI data centre investment
- [The Big Short](#) ⇨ star Michael Burry is betting against tech stocks - interesting
- AI industry has ~ \$1.2 trillion worth of debt ... more high-grade debt is now tied to AI than to US banks!
- Julien Garran (MacroStrategy Partnership) reckons [this is 4x bigger](#) ⇨ than the 2008 housing bubble
- If this debt get re-valued (e.g. because one of the AI giants goes bust) we may see a recession 4x worse than 2008