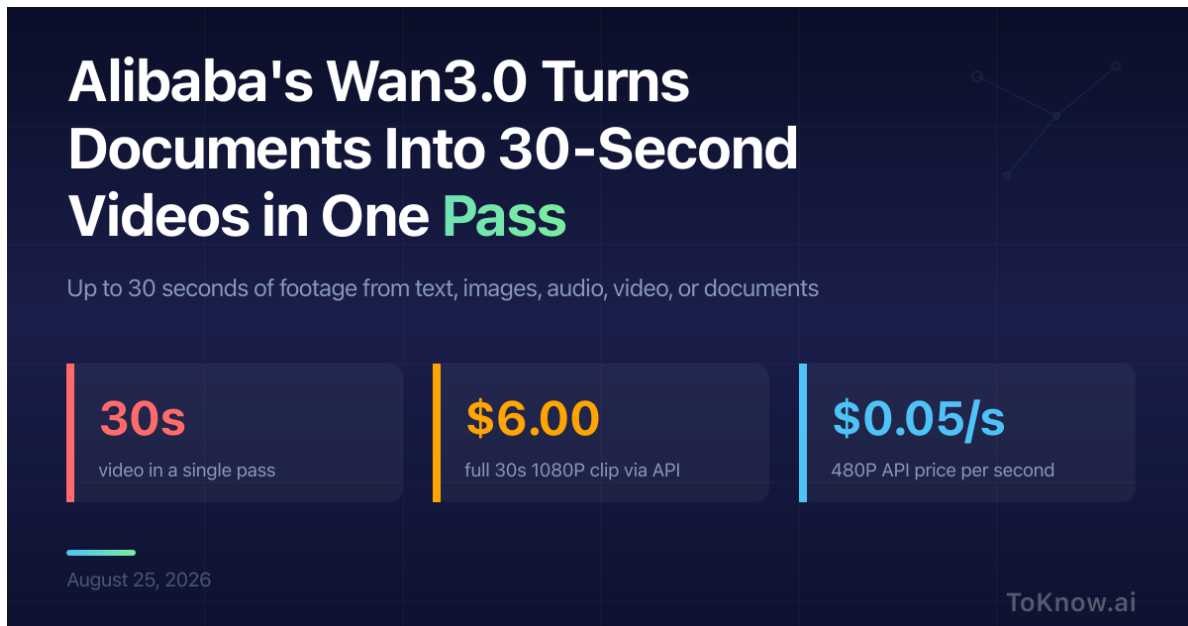


Alibaba's Wan3.0 Turns Documents Into 30-Second Videos in One Pass

Kabui, Charles

2026-08-25

[Read at ToKnow.ai](#)



The graphic features a dark blue background with a subtle grid pattern. In the top right corner, there is a faint, stylized line drawing of a person's head and shoulders. The main title is in large, bold, white and green font. Below the title, a subtitle in smaller white font reads 'Up to 30 seconds of footage from text, images, audio, video, or documents'. Three vertical bars in red, yellow, and blue separate three columns of information. The first column shows '30s' in red, with 'video in a single pass' below it. The second column shows '\$6.00' in yellow, with 'full 30s 1080P clip via API' below it. The third column shows '\$0.05/s' in blue, with '480P API price per second' below it. At the bottom left, there is a small green bar followed by the date 'August 25, 2026'. At the bottom right, the text 'ToKnow.ai' is visible.

Alibaba's Wan3.0 Turns Documents Into 30-Second Videos in One Pass

Up to 30 seconds of footage from text, images, audio, video, or documents

30s video in a single pass	\$6.00 full 30s 1080P clip via API	\$0.05/s 480P API price per second
--------------------------------------	--	--

August 25, 2026

ToKnow.ai

Alibaba officially [rolled out Wan3.0 on August 24](#), after a public beta that began August 6. It is a video generation model that produces up to 30 seconds of footage in a single pass, and the input side is the story: alongside text, images, audio, and video, it accepts [office documents](#), doc, xls, ppt, pdf, and markdown files, plus web page links, one file up to 100 MB or 50 pages per request. It runs on Alibaba Cloud Model Studio, with API pricing of \$0.05 per second

at 480P, \$0.10 at 720P, and \$0.20 at 1080P, so a full 30-second 1080P clip comes to \$6.00. Alibaba Cloud is [discounting API prices 30% until September 23](#).

A sales deck becomes a product launch ad, a training document becomes video courseware, a spreadsheet becomes animated charts, and a quarterly report becomes a narrated briefing, with no scriptwriter or editor in between. That collapses the usual multi-tool pipeline into one API call, and at \$1.50 per 30-second 480P draft, iteration is cheap. Wan3.0 also lets you edit instead of regenerate: the built-in editor can change visuals, plot, or dialogue on an existing clip. On-screen text rendering and audio texture are still weak spots.

Video generation is moving from short clips to anything-to-video, and when the source can be the documents people already make for work, generated video becomes the default way to present ordinary office content.

Read More: [ByteDance Seedance: one model that takes text, images, audio, and video together to make clips](#)

Sources:

- [Wan3.0: 30-Second AI Video Generation from Any Input \(Alibaba Cloud\)](#)
- [Alibaba launches Wan3.0 AI video model after \\$10 billion share sale \(Reuters\)](#)
- [Alibaba launches Wan3.0 video model with 30-second generation and document input \(TechNode\)](#)
- [Wan 3.0 AI Video Model: 30-Second One-Take, Document-to-Video \(PicLumen\)](#)

***Disclaimer:** For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. **Read more:** [/terms-of-service](#)*