

AI Video Policy Essentials

The **top 10 things** government communicators should know about deepfakes, AI disclosure, and disclosure, and staying out of the deposition. The back half you can use on Monday.

Dustin Haisler

Chief AI Officer & U.S. General Manager, Darwin AI

You wear three hats. Your policy probably covers one.

1

PUBLISHER

Your own AI-assisted video: avatars, dubbed translations, voiceover, PSAs. Your day job, and where your first violation comes from.

2

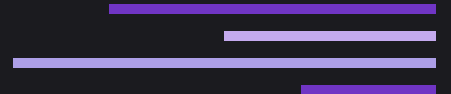
TARGET

A cloned voice of your city manager. Your seal
Your seal on a fake alert. A fabricated clip of
of your mayor, dropping Friday at 9pm.

3

VERIFIER

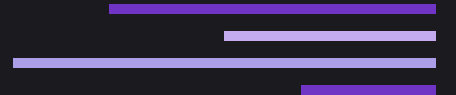
When a clip circulates, the public asks you if it's
you if it's real. You are the authority on your
your own content. If you can prove it.



Anyone can fake anyone. Including your mayor.

*Your nephew can do this. So can a scammer three time zones away.
Same afternoon.*

10



Consumer tools, official targets.



Minutes, not studios

A prompt in a consumer app returns broadcast-quality synthetic video. What took a VFX team in 2023 now takes a bored teenager.



Officials are the lure

The FBI keeps issuing advisories on AI impersonation of U.S. officials. In 2025, an AI voice clone of the Secretary of State contacted foreign ministers, a governor, and a member of Congress.



A video call is not ID

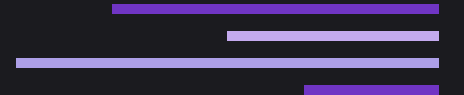
Finance, HR, and executive assistants are the top targets. 'I saw them on the call' no longer counts as verification.

The raw material is your own feed. Your leaders' voices live on your channels and your org chart is public. Every state introduced deepfake legislation in 2025 for a reason.

Government gets regulated first. first.

*The first wave of AI law isn't aimed at Big Tech. It's aimed at us.
Inventory it, disclose it, buy it right.*

09



The rules that already govern public-sector AI.



Inventory your AI

Federal agencies must publish annual AI use-case inventories (OMB). Connecticut and Texas passed state agency inventory laws, and the list grows every session. Know where AI touches your comms, including the unsanctioned tools.



Disclose your AI

Texas TRAIGA, live since January 2026: 2026: agencies must tell residents when when they're interacting with an AI system. This is the template other statehouses are copying.



Buy it with strings

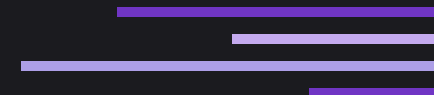
State AI procurement standards are landing now. Provenance support, disclosure terms, and audit rights belong in your next comms and media tool contracts.

Start with the inventory. If your agency can't list where AI touches its communications today, that's step one. Everything else builds on it.

The platforms already labeled
you.

*No legislature involved, no grace period. It's running on your official
accounts today.*

08



Platform rules beat state law to your feed.



TikTok labels you anyway

Integrated C2PA in January 2025; over 1.3 billion AI videos labeled since. Detection runs whether or not you disclose. What it catches unlabeled can be labeled for you, downranked, or pulled.



YouTube and Meta

YouTube requires the altered-content content toggle for cloned voices and and realistic synthetic depictions. Meta Meta unified the "AI Info" label across across Instagram and Facebook in February 2026, applied in-post.



The carve-out (same on all three)

AI captions, hashtags, descriptions, text overlays, and script help need no label anywhere. The label attaches to the media, not the planning around it.

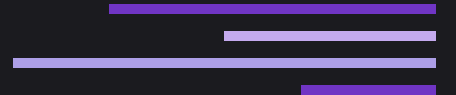
DO THIS

Match your agency's labeling trigger to the platform trigger. Don't invent a third definition your team has to translate at upload.

**Your mayor's voice
isn't yours to clone.**

*Even with good intentions. Consent first, in writing, with an expiration
expiration date.*

07

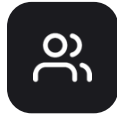


Likeness has lawyers now.



The ELVIS Act started it

Tennessee, 2024: the first law protecting a person's voice from AI cloning. A wave of state right-of-publicity updates followed, and the federal NO FAKES Act is still circling Congress. Voice and face are property now.



Yes, this means your own officials

Cloning your mayor's voice for Spanish Spanish dubs or building a presenter presenter avatar requires written consent, scoped to specific uses. And And decide now: what happens to the the avatar when they leave office?



And definitely not celebrities

No soundalike narrator voices, no style-of-a-star AI music under your videos. If the voice sounds like someone, someone owns it, and public agencies make attractive test cases.

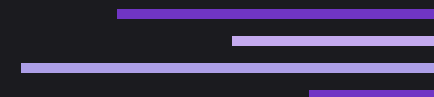
DO THIS

Add a one-page digital-replica consent form to your toolkit before your first AI dub. Signed once per official, revisited at term end.

“Made with AI” labels are becoming law.

Fifty states, fifty rulebooks. Only one binds your agency directly. Worth directly. Worth knowing which one.

06



Four mandates, and who they actually bind.

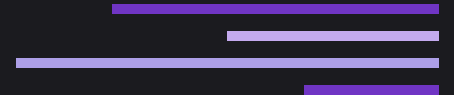
JAN 2026	Texas TRAIGA	Agencies must disclose when residents are interacting with an AI system. The only label law that binds government directly today.	GOVERNMENT
JUN 2026	New York S.8420-A	Conspicuous disclosure for ads using AI "synthetic performers." Reaches you if your agency's paid campaigns use AI talent.	ADVERTISERS
AUG 2026	California AB 853	AI providers must embed latent disclosures in AI video, audio, and images. Translation: your creative tools start labeling for you.	AI PROVIDERS
31 STATES	Election deepfake laws	Label-or-ban rules for synthetic political media, stress-tested by the 2026 midterms. Binds campaigns and distributors, not your official feed.	POLITICAL ADS

Net for agencies: only Texas binds government directly today. The rest reach you through your tools, your ad buys, and your residents. Direction of travel: label
Direction of travel: label everything.

**Your peers are already shipping
shipping. AI video.**

*Avatars, live translation, captions, first drafts. Policy exists to keep this,
keep this, not kill it.*

05



The productive side of the ledger.



Amarillo's "Emma"

A digital-human avatar that answers residents in 60+ languages, on video, around the clock. Built by a city of 200,000, not a federal lab.



Meetings in every language

Cities now live-translate council council meetings and press conferences with AI, multilingual multilingual access no translation translation bureau could staff.



The everyday stack

AI captions for WCAG compliance, Spanish dubs of PSAs, b-roll tagging, script first drafts. Google demos exactly this at 10:30 today.

Every one of these is legal and encouraged. Each needs a different level of disclosure, which is exactly where we're headed next.

The toolbox is rapidly expanding!

Avatars & presenters

Synthesia, HeyGen. A consistent on-camera presenter in 60+ languages, the Amarillo playbook. Consent paper first if it's a real official.

Voice & dubbing

ElevenLabs, Descript. Narration, Spanish dubs, fixing one flubbed word without a without a reshoot. Tier 3: disclose in the description.

Prompt-to-video

Google Veo (Flow), Nana Banana, Runway, Sora. B-roll you can't film: aerials, seasons, re-creations. Tier 1: label it on the face.

Inside tools you already own

Canva Magic Studio, Adobe Premiere + Firefly. AI baked into your existing licenses, and licenses, and both ship Content Credentials. Turn them on (see #2).

Repurpose & captions

Descript, Opus Clip. Turn an hour-long meeting into shorts, auto-caption everything for WCAG. Caution: CapCut is ByteDance, banned on government devices in a growing list of states.

Live translation

Wordly and similar. Real-time multilingual streams of council meetings and pressers, no interpreter bench required.

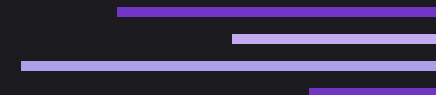
DO THIS

Pick from this menu with IT in the room. Check each tool for C2PA export and government terms, then write the winners into your approved-tools list (#4).

Your AI use needs a paper trail.
Your AI use needs a paper trail.
trail.

*If it isn't logged, it didn't happen. If it is logged, it's a public record.
record. Welcome to government.*

04



How to use AI without becoming the story.



Use approved tools, not personal accounts

Run AI through agency-sanctioned tools your IT team can see. Shadow AI on a personal login is how liability starts.



Keep a named human in the loop

Someone reviews and approves before publish, and the workflow shows who. 'The AI did it' is not a defense; it's a headline.



Log prompts and outputs

AI-assisted content is a public record in most states. Retention schedules apply to your prompts, drafts, and approvals.



Watch what you paste in

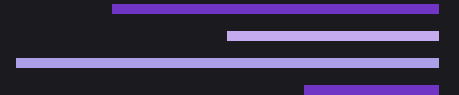
Resident PII, personnel matters, and non-public data never data never go into consumer AI tools. That's a breach, not breach, not a workflow.

The test: if a records request landed tomorrow, could you show what the AI did, what you changed, and who approved it?

**Label by tier,
not by vibes.**

Most policies say “disclose AI use” and leave the judgment call to whoever uploads on a Friday afternoon.

03



Four tiers. What actually trips a label.

TIER 1

Generated

A realistic synthetic depiction of a person, place, or event that did not happen.

Label on the asset, the post, and the archive.

TIER 2

Substantially altered

Real footage edited so a real person appears to say or do something they did not.

Label the same way. Or do not publish it.

TIER 3

AI-assisted

Dubbing, translation, synthetic voiceover of scripted copy, upscaling, cleanup.

Disclose in the description. Not on the face.

TIER 4

De minimis

Captions, hashtags, script drafting, text overlays, color correction, denoise.

No disclosure. Say so out loud in the policy.

Tier 4 matters as much as Tier 1. Stay silent on captions and color correction and your team labels everything, the label stops meaning anything, and you spent your credibility for nothing.

Steal this language. Seriously, it's why we wrote it.

VIDEO / POST CAPTION



"This video includes AI-generated content, reviewed and approved by [Agency] communications staff before publication."

WEBSITE FOOTER / PRIVACY POLICY



"[Agency] uses AI tools to help draft, edit, and produce produce some public communications, including social including social media and marketing content. Staff Staff review all AI-assisted content before publication. publication. Questions: [email]."

CHATBOT / AUTO-REPLY



"You're chatting with [Agency]'s virtual assistant, an automated AI service. To reach a person, reply HUMAN or call [number]."

AI-ENHANCED REAL FOOTAGE



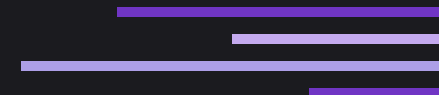
"Portions of this video were enhanced with AI tools tools (captions, audio cleanup). No content was materially altered."

Update the privacy policy once, reuse everywhere. Get legal's sign-off on the language one time, then it covers every post.

Detection is a coin flip.

You can't reliably catch a fake of your mayor. You can prove which video is the real one.

02



The math is not on detection's side.

63.3%

Human accuracy telling deepfake video from real, averaged across 22 studies. A coin flip with confidence.

3 sec

Of audio needed to clone a voice well enough to enough to fool a colleague on a call.

\$0

Cost of turning on Content Credentials (C2PA) in tools your team already licenses.

The asymmetry is the strategy: you can't reliably catch what comes in, but you can cryptographically sign everything you send out. Signed originals are what protect you when someone forges your principal.

DO THIS

Don't buy a detection product. Check the export settings in the video tools you already own and turn on C2PA signing this week.

The first sixty minutes of a deep fake.

A fabricated clip of your principal lands at 9pm on a Friday. This is a comms playbook, not an IT ticket.

0 - 15 MIN

Contain

Do not amplify. Do not quote-post the the fake. Capture it: URL, timestamp, timestamp, screen recording. Tell legal legal and IT. Confirm internally it's false false before you say it's false.

15 - 45 MIN

Counter

Publish on your verified owned channel first. Lead with what is true, never with the false claim. Link the signed original. One statement, not a thread.

45 MIN +

Close the loop

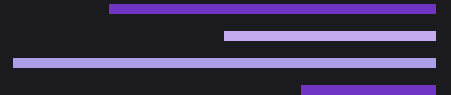
Platform reports, notify peer agencies agencies and your media list, log it for for the record. Then prebunk: tell people people where your real content lives and lives and how to check it.

Your only two advantages are a verified channel and a signed original. Both are built on an ordinary Tuesday, not on the Friday.

**Steal a policy.
Don't write one.**

Your agency doesn't need a novel. It needs a template and an afternoon.

01



Free places to start, today.



GovAI Coalition

Peer-built templates from 1,000+ agencies: AI policies, vendor agreement language, incident response. Free, government-made, battle-tested.



Darwin Policy Wizard

Answer a few questions, get a starter AI policy tailored to your agency. Free at policy.darwingov.com. Bring it to counsel, not a blank page.



What good policy covers

Approved tools. Disclosure rules (see #3). Records and retention (see #4). Human review. Incident response for the day the fake drops.

policy.darwingov.com · free starter AI policy in about 10 minutes. Yes, actually free.

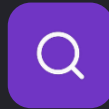
darwingov.com

If you can't see it,
you can't defend it.



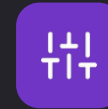
Visibility

Know every AI interaction happening across your agency, sanctioned or not.



Auditability

Prove what the AI did, what humans changed, and who approved it, on demand.



Enforcement

Your AI policy applied automatically, not filed in a drawer and forgotten.

That's what we build at Darwin. Visibility, auditability, and policy enforcement for government AI. Come say hi, or start with the free wizard: policy.darwingov.com

Dustin Haisler Chief AI Officer & U.S. GM, Darwin AI

Now go make great video.