

# 1. EXECUTIVE SUMMARY

Accessible video means any web-based video content that can be perceived and understood by people with disabilities, typically through features like synchronized captions for the hearing-impaired and audio descriptions for the blind. Recent advances in AI (speech recognition, natural language processing, computer vision) can automate much of this work. AI-driven tools can rapidly generate transcriptions and captions for video and even draft audio descriptions, greatly reducing time and cost. However, these automated outputs often contain errors or omissions, so human review remains essential to meet standards.

Complying with the Web Content Accessibility Guidelines (WCAG 2.1 AA) is now mandatory under U.S. law (for example, the 2024 ADA Title II rule requires state/local government content to meet WCAG 2.1 AA by April 2026, and Section 508 of the Rehabilitation Act references WCAG 2.0 AA). In New Jersey, the state has actively publicized these requirements, and courts are seeing a spike in ADA-related website accessibility suits (one NJ attorney filed ~60 suits in 2023 alone). NJ businesses are thus under pressure to ensure their digital content – including videos – is accessible. Federal and local regulations, along with high-profile litigation, make this a current priority: up to 20% of Americans are deaf or hard of hearing, and 96% of top websites have accessibility issues, so the stakes are high.

This report explains how AI-enabled workflows can produce WCAG-compliant videos. It defines key terms (e.g. captions, audio description, WCAG) and clarifies misconceptions (e.g. *automated captions alone do not guarantee compliance*). It reviews relevant laws: for NJ, Title II's new rule enforces WCAG 2.1 AA on public-sector video, while nationwide the ADA, Section 508, and FCC regulations mandate captions and descriptions for public video. We examine market factors: New Jersey firms are influenced by NYC/Philly markets and often operate in sectors like pharma, finance, construction, healthcare and education with strong accessibility needs.

We then delve into *how AI works in accessible video production*. Speech-to-text engines (e.g. Google, AWS, IBM) create raw captions or transcripts; specialized AI can propose audio descriptions; and tools can analyze videos for color contrast or scene content. Real-world workflows typically use a hybrid approach: automated output plus human editing. Examples of use cases include NJ construction companies training staff with captioned videos, NJ hospitals

providing patient education with descriptions, universities offering lecture videos to online students, and any corporation ensuring compliance training videos are accessible.

Cost and resource analysis shows captioning (human-reviewed) usually runs on the order of \$1–15 per video minute, while professional audio description can be \$15–75 per minute. Local NJ factors (e.g. unionized talent rates, travel, taxation) can push costs toward the higher end of these ranges. Turnaround times vary: a captioned short video can be done in 1–3 days, whereas audio description often takes 1–3+ weeks due to scripting and recording. We provide a comparison table of methods (AI-generated vs. human, etc.) to highlight tradeoffs.

Finally, we outline benefits vs. limitations: AI greatly speeds up routine tasks, but cannot yet replace human judgment for quality and nuance. Risks include assuming “*auto-captions are enough*” (they do not meet WCAG by themselves) or overlooking non-speech information (sound effects, speaker IDs). We warn of common mistakes (e.g. posting videos without captions, relying on burned-in text, skipping QA) and discuss failure modes (e.g. captioning errors, ADA lawsuits). Alternative solutions (fully manual captioning, transcripts-only, sign-language overlays) are compared against AI-driven methods. The report concludes with future outlook (AI improvements, possible new regulations by 2028) and checklists for decision-makers and vendors – covering both NJ-specific considerations and U.S.-wide best practices.

## 2. DEFINITIONS & TERMINOLOGY

- **Web Accessibility:** Designing digital content so people with disabilities can use it. According to the W3C, “web accessibility means that websites, tools, and technologies are designed and developed so that people with disabilities can use them,” specifically that they can *perceive, understand, navigate, and interact with the Web*. This encompasses visual, auditory, cognitive, and motor disabilities. Good accessibility also benefits non-disabled users (e.g. captions in noisy environments, transcripts for quick reading, etc.).
- **WCAG (Web Content Accessibility Guidelines):** A set of globally recognized technical standards from the W3C for making web content accessible. WCAG 2.1 (the current version) at Level AA is widely adopted as the legal and practical benchmark in the U.S.

(per DOJ's Title II rule and Section 508). It has four core principles (POUR: Perceivable, Operable, Understandable, Robust) and about 80 success criteria (e.g. 1.2.2, 1.2.3, 1.2.4) covering media such as video. WCAG provides the criteria that captions, descriptions, and related features must meet (e.g. captions must be accurate, complete, synchronized). Failure to meet WCAG typically means non-compliance with ADA/Section 508.

- Accessible Video: Video content (live or recorded) that includes the necessary features so people with disabilities can fully understand it. Key components include:
  - Captions: Text equivalents of all spoken dialogue and meaningful audio in the video, time-aligned with the media. Captions allow deaf or hard-of-hearing users (and many others) to follow the content. They should include speaker identification and non-speech sounds (e.g. “[phone ringing]”). Closed captions can be toggled on/off; open captions are burned in and always visible (though WCAG prefers closed captions).
  - Transcripts: A full text version of the video’s dialogue and narration. Unlike captions, transcripts are not time-synced and usually are a separate text file. They are useful for users who cannot access the video player or need a quick text overview. Per Section 508 guidance, a transcript is “the output of a process in which speech or audio is converted into a written, plain text document”. Transcripts *complement* captions but do not replace them in meeting WCAG (except in very limited cases like video-only media).
  - Audio Descriptions: Narrated explanations of essential visual information (actions, scene changes, text on screen) inserted into pauses in the main audio track, for the benefit of blind or low-vision users. AD provides “access to visual information by means of a verbal representation of visual elements”. For example, describing what’s happening on screen when people on screen are not speaking. (If a video’s essential info is already conveyed by dialogue, separate AD may not be needed.)
  - Assistive Technologies: Tools used by people with disabilities to access content. For video, this commonly includes screen readers (reading on-screen

text), braille displays, voice interfaces, and media players that support caption toggling. Ensuring accessible video means these tools can do their jobs: e.g. screen readers should be able to announce caption availability, and players should allow increasing text size and contrast.

- ADA (Americans with Disabilities Act): A U.S. civil rights law prohibiting discrimination on the basis of disability. *Title II* covers state and local government services, and *Title III* covers “public accommodations” (private businesses open to the public). While the ADA itself does not list specific technical standards, DOJ rules and case law have adopted WCAG as the standard for web accessibility. In 2024, DOJ’s final Title II rule explicitly requires WCAG 2.1 AA for government web/app content. (Title III litigation often also uses WCAG as the benchmark.)
- Section 508: A U.S. federal law (Rehabilitation Act) that requires federal agencies (and vendors they work with) to make their electronic and information technology accessible. The Section 508 Refresh (2017) aligns with WCAG 2.0 Level AA as the standard for web content. So federal websites and any entity handling federal information must meet comparable criteria (e.g. captions on all multimedia).
- Captions vs. Subtitles: These terms are often confused. *Captions* (especially “SDH” subtitles for deaf/hard-of-hearing) include not only dialogue but non-speech audio cues and speaker IDs, for those who cannot hear the audio. *Subtitles* (non-SDH) typically only translate dialogue into another language for hearing viewers. Subtitles for translation do not meet accessibility needs (they omit sound effects, etc.). For WCAG compliance, captions (with non-speech cues) are required – mere foreign-language subtitles are not acceptable.
- Common Misconceptions:
  - “Auto-captions are sufficient.” In fact, AI-generated captions often contain errors (misheard words, missing sound cues) and therefore do *not* meet WCAG criteria by themselves. They must be reviewed and corrected.
  - “Video transcripts alone equal compliance.” Transcripts help but WCAG’s synchronized media criteria generally require captions on the media itself;

transcripts do not fulfill the caption requirement unless the video is an alternative to text (rare).

- “Only large organizations need accessible video.” Actually, ADA and other laws apply broadly (including small businesses if they are public accommodations), and inclusivity and SEO benefits apply at any scale.

### **3. REGULATORY & COMPLIANCE LANDSCAPE**

#### **A) NEW JERSEY LENS**

New Jersey has proactively implemented the federal Title II ADA rule for state/local governments. Effective June 2024, NJ agencies must follow DOJ’s rule requiring WCAG 2.1 AA compliance on all web and app content (new and existing) by April 24, 2026. This means NJ state and local websites (including videos) must have accessible features (e.g. captions, audio descriptions) at WCAG 2.1 AA standards. NJ emphasizes accessibility in procurement: agencies are expected to identify and update inaccessible content and to archive or remediate older digital materials that do not comply. NJ’s statewide guidelines explicitly cite the WCAG 2.1 Level AA standard and provide resources (like ADA.gov fact sheets) for compliance.

There is also growing legal attention in NJ: private plaintiffs are increasingly suing businesses for inaccessible online content. The New Jersey Civil Justice Institute notes a surge in ADA website suits filed by visually-impaired plaintiffs, including dozens of cases in 2023 targeting major corporations. While most reported cases involve websites, the same legal theory could apply to inaccessible videos. In practice, NJ organizations (especially public entities and large companies) are sensitive to such risks. Public-sector clients (universities, municipalities) must meet the ADA’s “effective communication” mandate; for example, a captioned video is a way to provide “effective communication” for someone with hearing loss. In procurement, many NJ agencies now include explicit WCAG requirements in RFPs and contracts, though no single NJ-specific video law exists beyond the federal standards. (NJ does not have its own digital accessibility law beyond adopting ADA/Section 508 rules.)

#### **B) NATIONAL (USA) LENS**

Federal Laws and Standards: At the national level, the Department of Justice enforces ADA Title II (state/local governments) and Title III (public accommodations). In March 2024, DOJ finalized a rule under Title II setting WCAG 2.1 AA as the technical standard for all state/local public web content. Section 508 (federal agencies) already requires WCAG 2.0 AA (similar to 2.1 AA) for their online media. The FCC's 21st Century Communications and Video Accessibility Act (CVAA) also mandates captions for online video content, continuing the requirement that closed captions be provided for all publicly-posted audiovisual media. In practice, most U.S. jurisdictions now treat WCAG Level AA as the benchmark for compliance, whether under ADA, Section 508, or other regulations.

Enforcement Trends: Across the U.S., ADA website lawsuits have risen sharply (e.g. thousands filed annually) because there was no definitive rule for web content. New DOJ leadership is now signaling stricter enforcement. At the same time, DOJ has acknowledged that AI captioning technology is evolving, so it *encourages* best practices rather than dictating a fixed accuracy level. Notably, W3C guidance warns that "current auto-captioning technology would not allow your agency to meet the minimum standard for pre-recorded media", reflecting national concern that auto-only solutions fall short of WCAG requirements.

Industry Guidelines: Numerous agencies and industry groups provide guidance. The FCC (and DOJ) require *accurate, complete, synchronized captions*. W3C's WCAG SC 1.2.2–1.2.5 detail media requirements (closed captions for pre-recorded and live media, audio descriptions, etc.). Accessibility firms note that while AI can bootstrap the process, human review is the only method that truly meets WCAG criteria. For example, Level Access notes that only human-edited captions (often called "closed captions" in WCAG terms) are considered fully compliant. Both public and private sectors are advised to plan for WCAG compliance proactively, not reactively (fines and lawsuits can be far costlier than implementing captions early).

## **4. MARKET & INDUSTRY CONTEXT**

### **A) NEW JERSEY LENS**

Sector Mix and Buyer Behavior: New Jersey's economy spans healthcare, pharmaceuticals (centers like J&J, Merck), finance (Jersey City, Newark), manufacturing, construction (Bridges,

highways), higher education (Rutgers, Princeton), and professional services. Many NJ enterprises serve large, diverse markets, including NYC and Philadelphia nearby. This proximity means NJ organizations often compete or collaborate with NY/PA firms, adopting similar tech and standards. For video content, NJ companies increasingly recognize that accessibility can be a differentiator and risk mitigator – especially after high-profile lawsuits and the ADA rule announcements. Even smaller NJ businesses now seek multimedia vendors or agencies that guarantee WCAG-compliant outputs, and larger firms often embed accessibility criteria into marketing and training workflows.

**Influence of NYC/Philly:** Being adjacent to two major urban centers, NJ firms benefit from (and must match) the media production capabilities of NYC/Philly. For example, NJ's film and digital media industry is growing rapidly, fueled by state tax incentives and new soundstages. This infrastructure also supports corporate and educational video production. However, it also means local businesses are aware of nearby accessibility enforcement: New York State has its own accessibility laws, and Philadelphia-area institutions likewise emphasize ADA compliance. NJ organizations often draw on talent and standards from these areas, so trends in NYC/PA (such as using hybrid AI-captioning services) spill over into NJ practices.

**High-Risk Sectors:** In NJ, sectors like healthcare (hospitals, insurance) and education (colleges, K-12) are high-profile for ADA compliance, since they provide critical public services. For example, a hospital's patient-education videos or insurance training modules must be captioned. Public services (NJ Transit, government agencies) also must meet Title II standards. In manufacturing and construction, safety training often uses video; these industries must caption content to avoid excluding any workers or running afoul of OSHA-related communication requirements. Finance and tech firms (e.g. banks in Jersey City, fintech companies) increasingly include accessibility in corporate social responsibility and compliance programs – partly to serve all customers, partly to avoid class-action risks seen in other states.

**Procurement Patterns:** New Jersey state and municipal agencies follow the Title II rule by issuing RFPs that require vendors' digital materials to meet WCAG 2.1 AA. Large NJ corporations, particularly if they serve public contracts (e.g. contractor on a government project), often mimic this by adding accessibility stipulations in vendor contracts. Media

production agencies in NJ are aware of these requirements, so many now offer “accessible video” as a service. Local trends include combining in-house marketing teams with specialized vendors; for instance, an NJ university might transcribe lectures using campus tools, then outsource final captioning to a compliance firm. Union labor can play a role – NJ, as part of the NYC media zone, generally pays union rates to on-site camera and sound crews, which increases the baseline cost of video projects (affecting budget calculations for adding accessibility).

## **B) NATIONAL (USA) LENS**

**U.S.-Wide Trends:** Across the country, there is growing corporate investment in accessible multimedia. Surveys indicate that a vast majority of companies recognize the *need* for accessibility, though many are still behind on implementation. Major industries (tech, finance, healthcare) often lead because of either regulatory oversight (e.g. federal funding triggers Section 508) or consumer expectations. Enterprises are more likely than SMBs to have formal accessibility programs; small businesses sometimes start addressing video accessibility only after an accommodation request or lawsuit.

**Regional Variation:** While the technical requirements (WCAG AA) are national, implementation can vary by region. Coastal states (CA, NY, FL, NJ) have seen more litigation, pushing local businesses to be proactive. In contrast, some heartland regions have had fewer high-profile suits, so business compliance may lag slightly. That said, cloud-based tools and services make accessibility work broadly consistent: a company in Texas can use the same AI captioning platform as one in New Jersey. For nationwide organizations, standards are typically set at the corporate level to meet the strictest applicable rules (often those for Title II or Title III ADA).

**Industry Benchmarks:** Nationally, accessibility teams compare projects by cost and quality. For example, a Fortune 500 company might set a corporate policy like “no video goes live without 99% accurate captions,” aligning with industry norms. Many U.S. organizations plan captioning budgets roughly 1–2% of their video production costs (as a rule of thumb), and expect turnaround times on the order of 24–72 hours for captioned content. Industry whitepapers note that 80% of caption users aren’t deaf (they watch videos in sound-off environments), highlighting the broad value of captions.

## 5. HOW AI ENABLES ACCESSIBLE VIDEO (TECHNICAL / OPERATIONAL BREAKDOWN)

AI technologies can automate several key steps in making video accessible:

- **Speech-to-Text and Captioning:** Modern automated speech recognition (ASR) services (e.g. Google Cloud Speech-to-Text, Amazon Transcribe, IBM Watson) can ingest video audio and output time-synced captions or transcripts. These AI-generated captions capture the spoken words rapidly. In practice, most workflows use *AI plus human review*: the AI produces a draft caption file, and then a human editor corrects errors, adds speaker labels, and notes non-speech sounds. This hybrid approach dramatically cuts effort: raw ASR may have 70–90% accuracy, but with editing it reaches near-100%. A typical NJ vendor might run a 30-minute video through ASR (minutes of processing) then spend 30–60 minutes reviewing it.
- **Machine Transcription:** In addition to captions, AI can generate full video transcripts (essentially the same output as captions, but in paragraph form). Some organizations use these for SEO or to enable search. Advanced NLP (natural language processing) tools can even produce summary transcripts or chapter markers. Though WCAG doesn't require transcripts *instead* of captions, having both (provided together) is a strong practice.
- **Automated Audio Description (AD):** AI for AD is emerging. Some vendors (e.g. Verbit, 3Play Media) use machine learning to pre-generate scripts describing visual scenes, which human describers then refine. AI can help spot scene changes or recognize key objects (via computer vision models), suggesting what to describe. For example, an AI might flag a slide change or an action for potential narration. However, fully automated AD (from start to finish) is not yet reliable enough to meet WCAG on its own, because it can misinterpret context or tone. Instead, the typical production uses AI to speed up parts of the AD workflow – for example, by auto-transcribing the existing audio to identify quiet moments, or by using voice synthesis to generate a draft narration under human supervision.

- Visual Contrast and Readability Checks: Some AI or rule-based tools (e.g. WAVE, WebAIM Contrast Checker) analyze video frames for text contrast, font size, color-blind friendliness, etc. For instance, if a video has on-screen text (titles, bullets), an AI could check the RGB contrast ratios against WCAG minima. In practice, this is usually a smaller part of video workflows: designers often manually ensure high contrast in graphics (NJ branding guidelines even reference 4.5:1 ratio). However, tools can automate checking and flagging any violations.
- Scene Detection & Tagging: Advanced computer vision can decompose video into scenes or identify objects/people (face recognition is less relevant ethically, but object recognition can tag settings). This is more common in large-scale media libraries (for search) than in typical corporate videos. In accessibility, scene detection might help generate an initial list of visuals to describe or categorize video chapters. For example, if an AI detects “kitchen scene” followed by “office scene”, it could hint to the creator about necessary descriptions.
- Language Translation and Localization: For videos serving multilingual audiences, AI can auto-translate captions into other languages. Systems like Google Translate or Microsoft Translator can create subtitles (or alternative audio tracks) in dozens of languages. This is adjacent to WCAG but can improve global accessibility. However, translated captions also require review for accuracy. AI helps speed this; e.g. a NJ-based firm using English captions could deploy a cloud API to generate Spanish subtitles, then have a bilingual editor refine them.
- QA and Compliance Validation: After content is prepared, AI-based accessibility testing tools (e.g. axe Auditor, HTML validation for web, or custom scripts) can verify caption file formatting and check basic criteria (synchronization, presence of captions). Human testers with disabilities should also review the final product. The operational workflow typically looks like this: content creator exports a video file → ASR engine creates captions → human editor refines captions and transcripts → (optionally) AD script is generated and reviewed → final captions and AD are embedded in a compliant player or distribution format → compliance specialists run QC tests (caption

completeness, error rate, color contrast) → video is published with accessible features.

## **A) NEW JERSEY LENS**

In New Jersey organizations, these steps often occur within a tightly managed project. A typical workflow in a large NJ company might involve: the *marketing or training team* producing a video, then handing it off to an *accessibility/compliance team* or external vendor for remediation. Local AV/IT vendors (e.g. firms in Newark or Jersey City) commonly offer “AI captioning + editing” as a package. Internal stakeholders (e.g. legal and communications) review the captions/descriptions before final approval, since mistakes could pose legal risk. Turnarounds are influenced by NJ calendars (e.g. local holidays, union rules); for instance, if relying on union voice talent for audio description, scheduling can add extra days. Procurement processes (especially for government or education clients in NJ) may require competitive bidding to select the AI-accessibility vendor, which adds lead time. Overall, NJ companies are increasingly integrating AI tools directly: for example, a NJ college’s media department might use Adobe Premiere’s built-in speech recognition to auto-generate a caption track and then refine it with a human team.

## **B) NATIONAL (USA) LENS**

Nationwide, scalable models have emerged. Enterprise organizations often use cloud APIs (Amazon, Google, Azure) directly via automated pipelines: e.g. a corporation’s intranet could automatically caption any uploaded training video. SMBs may rely on turnkey services (e.g. Rev.com, 3Play) that combine AI and human work. Across states, the technical approach is similar, but speed and scale differ: large US firms might caption hundreds of hours per month, whereas a small shop might only update a few videos annually. In general: speech-to-text AI runs quickly everywhere, so initial processing is fast. Localization of teams varies: some companies have in-house accessibility experts; others contract out-of-state. The notable difference is regulatory compliance: in the federal sphere, workflows might be audited under Section 508 standards, whereas private companies follow internal policies. The bottom line is that AI-enabled workflows allow accessible video production to scale – but at the national level, organizations of different sizes implement it to varying extents. For example, a nationwide

retailer might have a centralized captioning workflow for all marketing videos (leveraging AI to meet WCAG demands), whereas a small business in a less regulated state might only caption videos reactively.

## **6. BUSINESS USE CASES (PRACTICAL APPLICATION)**

Approach: We list example scenarios by industry and company size. Each scenario illustrates how a New Jersey organization might use AI-powered tools to meet WCAG for video content.

### **A) NEW JERSEY LENS (Examples)**

1. **Construction Company Safety Training (Mid-market):** A New Jersey construction firm produces training videos (e.g. “Worksite Safety Procedures”) for on-site crews. To ensure WCAG compliance, they use an AI captioning service: the raw captions from ASR are reviewed by a compliance staffer to add speaker IDs and sound cues (e.g. “[Jackhammer operating]”). They also generate a transcript with the same content. The final captioned video is used on their LMS (learning management system). If the state inspector audits their training docs, the company has accessible content ready.
2. **Manufacturing/Energy Compliance Video (Enterprise):** A NJ chemical manufacturer (subject to OSHA regs) releases an emergency-response video. They employ an AI tool to auto-translate the script and generate multilingual captions (English and Spanish). Human linguists then refine those captions. This means Spanish-speaking workers (NJ has large Spanish-speaking populations) can follow safety information. They also commission an AI-assisted audio description track explaining on-screen hazard symbols, which a professional describer quickly polishes.
3. **Healthcare Patient Information (Large Hospital):** A New Jersey hospital creates patient-education videos (e.g. “Understanding MRI Results”). They upload recordings to a secure system that runs voice-to-text for captions. Healthcare compliance staff review them (ensuring medical terms are correct). For blind patients, the hospital’s media team uses an AI voice synthesizer to draft audio descriptions of essential visuals (charts, diagrams) – then a staff member with a sight impairment reviews and finalizes the narration. The result is a WCAG-compliant video on the hospital’s portal.

4. Higher Education Lecture (University): A NJ public university records classroom lectures. Using campus accessibility resources, the video is passed through an AI transcription service; student workers then format the output as captions and transcripts. Because many lecture videos are long, this saves weeks of work. The university also experiments with AI tools that identify slide text and convert it to on-screen captions. Professors review the content before it's posted to Blackboard, ensuring compliance. This makes the university's distance-learning program accessible to deaf/hard-of-hearing students without delaying content release.
5. Corporate Compliance Training (Finance): A Jersey City bank produces mandated compliance training (e.g. "Anti-Fraud Policies"). They have a multimedia agency that uses AI captioning (paired with human QA) so employees in noisy offices or with hearing impairments can understand the material. A law firm consultant checks the captions against regulations. The bank also uses AI for keyword search: captions feed into their content management system so staff can search video transcripts for policy terms.
6. Logistics / Transportation (NJ Port Authority): The Port Authority of NY/NJ releases public-service videos (e.g. "How to Navigate Newark Airport"). These videos are captioned automatically via AI, then overseen by a digital accessibility specialist to meet WCAG. The videos also include audio descriptions so visually-impaired travelers get important visual information (e.g. "Passengers offloading baggage"). Because the Port Authority serves an international population, they also auto-translate captions into multiple languages for displays.
7. Technology Company Demo (SMB): A small NJ software startup creates marketing demos showing their new app. To ensure they don't inadvertently exclude any user, they use the captions feature built into their video editor (an AI-based tool) and then hire a freelancer to proofread. They don't provide full audio description (since the app demo is mostly voice-driven), but they do caption all dialogue and on-screen text. This makes the startup's online videos ADA-friendly and SEO-friendly (captions improve searchability).

## **B) NATIONAL (USA) LENS (Examples)**

1. Distance Learning Platform (Education): Across the U.S., ed-tech companies use AI captioning to make online courses accessible. Automated transcripts are quickly generated for each lesson, then refined by staff. Audio description is added for important visuals (graphs, experiments) either by AI-aided scripts or human narrators, depending on budget. For instance, a nationwide MOOC platform might caption in multiple languages using automated translation, then verify quality in key courses.
2. Streaming Media (Entertainment): Large streaming services employ AI to create initial captions for thousands of hours of content. While Hollywood TV shows often use human captioners, tech and indie producers rely on machine algorithms (with spot checks). In rare cases where budget allows, AI is also used to draft audio descriptions, but typically professional describers ensure film/TV compliance with laws like the CVAA.
3. Corporate Communications (Various Industries): Enterprises nationwide produce corporate videos (annual reports, internal town halls, training). Many use enterprise accessibility platforms (often cloud-based) that auto-caption any uploaded video. An internal accessibility team performs quality checks. Audio description might be outsourced via specialized services when budgets allow. Global companies often adapt this by auto-translating English captions into other languages with AI, supporting accessibility for international employees.
4. Government and Public Announcements: U.S. state and federal agencies routinely caption all public videos (press briefings, announcements). With limited budgets, they increasingly rely on automated captioning supplemented by manual corrections. After the 2024 Title II rule, even more local governments plan to integrate AI caption workflows. The VA and other federal entities follow Section 508, often using vendor platforms that combine AI and manual work.
5. Healthcare Nationwide: Similar to NJ examples, major hospital networks across the U.S. use AI tools for patient-facing videos (e.g. COVID-19 information). Federally-funded institutions must have accessible content; AI helps keep up with large

volumes (hundreds of hours of video) by automating initial caption creation.

Multi-state providers may have in-house accessibility teams to finalize this content.

6. SMB Marketing (Nationwide): Even small businesses in other states use consumer tools: YouTube’s auto-caption feature (requiring post-editing) or cloud services offering cheap AI captions. For example, a small retailer in California might use an online tool to caption its product videos, then have an employee quickly check them. While not perfectly WCAG-compliant initially, these SMBs are trending toward better accessibility as awareness spreads.

These use cases illustrate that AI-augmented workflows – from education to healthcare to corporate – are enabling accessible video production at scale, both in New Jersey and nationwide. The specific applications (training, marketing, education) may differ, but the underlying tasks (captioning, describing, translating, QA) are common.

## 7. COSTS, RESOURCES & TIMEFRAMES

Producing accessible videos incurs extra costs (and time) beyond standard production. Below is a summary table and discussion of typical ranges:

<b>Service/Workflow</b>	<b>Cost per Minute (Approx.)</b>	<b>Turnaround Time</b>	<b>Comments</b>
AI-augmented Captioning (AI + human review)	\$1–5	Hours–1 day (30-min video)	Fastest, uses machine transcripts + edits; includes speaker IDs & sounds; meets WCAG after review.
Manual Captioning (professional)	\$5–15	1–3 business days	Highest accuracy (>99%); covers all audio; most expensive.
Audio Description (human)	\$15–75	1–3+ weeks	Human scripts & voiceover; union talent drives cost; premium service.

AI-assisted Audio Description	\$10–30 (estimate)	1+ weeks	Uses AI to draft script; human ensures quality; faster/cheaper than fully manual.
Translation/Multilingual Captions	\$10–20 per language (human-reviewed)	1–5 days per 30-min video	Machine-translate plus native editor; additional cost per language.

*Notes:* As the table shows, captioning is relatively affordable: many vendors price from around \$1/min (for basic AI-assisted captioning) up to \$15/min for top-tier human transcription. A 10-minute training video might cost \$10–\$150 to caption, depending on service level. Audio description remains significantly more expensive due to skilled labor. For example, a single 5-minute instructional video could cost from \$75 to \$375 just for descriptive narration.

New Jersey considerations: NJ falls in the higher-cost tier for U.S. production, largely due to its media industry’s proximity to NYC. Union contracts (SAG-AFTRA for voice talent, IATSE for production crews) often apply in NJ (especially north Jersey projects) – this can add 20–30% to voiceover rates. Travel and logistics (e.g. hiring an NYC actor or studio) may appear on NJ budgets where other states use local non-union talent. On the procurement side, NJ public agencies sometimes bundle multimedia services into larger contracts, which can lower per-minute costs through volume discounts. But individual businesses in NJ should budget accordingly: vendors serving the NJ market typically align their pricing with NYC-level rates.

National context: Across the U.S., captioning costs are broadly similar (the \$1–15/min range). However, some variation exists: major markets (NY, LA, SF) charge at the upper end; smaller markets or overseas companies can offer lower rates. Turnaround times also differ: enterprise plans (like 3Play’s “Pro/Enterprise”) may promise same-day turnaround on large batches, whereas small vendors may need a week. Generally, AI speeds up turnaround – many vendors now offer caption drafts within minutes of upload. For audio description, no comparable automation currently yields same-day delivery – expect weeks due to multi-step production.

Cost Drivers: Key factors driving these costs include:

- Video length and complexity: Longer videos cost more linearly. Highly technical language or multiple speakers increase transcription time.

- Accuracy requirements: Demanding 99% accuracy (for WCAG) means human editing, raising cost. Automated captions alone are cheaper but *not* WCAG-compliant.
- Quality level: Providers offering “broadcast-quality” captioning (with full verbatim scripts and spot-on timing) charge more than basic services.
- Union rates: As noted, location matters. For NJ, using unionized talent (common for professional narration) inflates AD costs.
- Turnaround time: Faster deadlines incur rush fees. A 24-hour turnaround will cost more than a standard 3–5 day schedule.
- Volume discounts: Large organizations captioning thousands of minutes annually will pay less per minute (often below \$1/min on enterprise contracts) than a one-off small project.

Example: According to 3Play Media, a fully in-house captioning workflow (no AI) for 1 hour of video might take 12 hours of labor (\$165 total at \$15/hr). Outsourcing that at \$3/min would cost \$180 (\$3×60) with minimal internal time. Audio description is a step change higher: even with AI aids, professional voices and scripts mean ~\$20–30 per minute is common, so a 1-minute narrated description can approach the cost of a 5-minute caption segment.

## **8. BENEFITS VS LIMITATIONS**

### **A) NEW JERSEY LENS**

Benefits (NJ): Accessible videos extend reach to NJ’s diverse population (multilingual and disabled audiences) and improve employee training effectiveness. For example, captions benefit non-English speakers and people with hearing loss (NJ has sizable deaf/hard-of-hearing communities), and also commuters who watch videos in noisy NJ transit or quiet libraries. Captioned NJ business videos can boost SEO and user engagement; studies find 80% of caption users aren’t deaf, indicating broad utility. In education and healthcare, accessibility enhances inclusivity and can improve outcomes (e.g. blind patients understanding scan results via AD, or students with learning disabilities benefiting from captions). Legally, NJ companies gain compliance peace-of-mind: having accessible videos can prevent costly litigation. A recent

3Play survey highlights that proactive captioning avoids settlement costs (e.g. Netflix, FedEx cases).

Limitations/Trade-offs (NJ): Using AI has limits. Current auto-caption engines may mis-transcribe New Jersey-specific terms or names (e.g. local geography), requiring careful QA. AI often misses context cues (music, ambient sounds) that humans catch. Audio description via AI is even less mature – automated scripts can seem unnatural or omit subtle cues. Time and budget constraints can also be tighter for smaller NJ businesses, meaning some may delay full compliance. There is also a learning curve: NJ organizations must invest in training or hiring staff who understand WCAG nuances. In tightly regulated sectors like finance, the bar for “good enough” is high, so relying on AI alone can leave gaps: as one guide notes, AI needs “human assistance to create closed captions” that actually meet standards.

## **B) NATIONAL (USA) LENS**

Benefits (USA): Across the U.S., AI-enabled accessibility has similar advantages. Businesses that caption all video not only serve customers with disabilities but also improve overall viewer comprehension and engagement. From a practical standpoint, automated tools drastically reduce labor: what once took hours of transcription can now be minutes, enabling real-time captioning of live or rapidly-produced content. Compliance-wise, using AI workflows with checks helps organizations meet WCAG 2.1 obligations more efficiently. Especially for large multi-state companies, a unified AI-driven process ensures consistency nationwide. Industry data show that accessible content often results in positive brand perception and can drive innovation (e.g. voice interfaces developed for accessibility have wider product uses).

Limitations (USA): The core limitations mirror the NJ case: AI accuracy is not 100%. For example, Level Access points out that AI-generated captions “often inaccurate and incomplete” without human checks. In the U.S., reliance on out-of-the-box solutions (like YouTube’s auto-captions) can lead to non-compliance. Another challenge is over-reliance on technology: an organization might assume “having captions” is done by a quick AI pass, but regulators expect precise adherence to WCAG criteria (complete, synchronized, with all audio info). For audio description, the limitation is even more pronounced: without skilled describers, AI audio narration can fall flat on creative nuance. Finally, video content varies wildly; for example,

fast-paced entertainment or highly technical videos may confound AI more than typical corporate dialogue, necessitating heavy manual corrections.

Overall, AI is a powerful enabler but not a silver bullet. The “sweet spot” is using AI where it excels (speech recognition, initial caption drafts) while retaining human oversight for quality and legal compliance. This balanced approach yields the benefits (speed, scalability) without the critical failures.

## 9. RISKS, FAILURE MODES & COMMON MISTAKES

### A) NEW JERSEY LENS

- **Over-Reliance on Auto-Captions:** A NJ firm might be tempted to use only the automatic captions provided by hosting platforms (YouTube, Vimeo, etc.). This is risky: as Section 508 guidance notes, *“auto-captioning technology... would not allow your agency to meet the minimum standard”*. In other words, an NJ company solely relying on AI without review risks non-compliance.
- **Incomplete or Inaccurate Captions:** Common mistakes include missing speaker changes, failing to caption music or sound effects, or leaving out text that appears onscreen. WCAG success criterion 1.2.2 requires *all* spoken words and sound cues to be captioned; omitting these (e.g. no “[phone ringing]” annotation) is a failure. Unverified AI output often has such gaps.
- **Neglecting Audio Descriptions:** Some NJ organizations view AD as optional. Failing to provide descriptions for critical visuals can violate WCAG 1.2.3 (especially for video-only content). For instance, a promotional video showing charts without narration might be inaccessible to blind users without AD. The risk here is accessibility oversight: NJ laws don’t explicitly mandate AD in every case, but WCAG/DOJ expect it “when the visual info is not otherwise conveyed”.
- **Ignoring Mobile or Platform Issues:** If a NJ business posts videos to platforms (Instagram, Twitter) that auto-play muted, they might omit captions. This excludes deaf users and those who rely on captions in silent autoplay contexts.

- Lack of Testing: Failing to test with real users with disabilities is a mistake. NJ teams sometimes skip actual user testing, assuming good enough. This can miss issues (e.g. timing errors) that technically pass automated checks but trip real users.

## **B) NATIONAL (USA) LENS**

- Legal Liability: Across the U.S., the biggest risk is ADA lawsuits or OCR complaints. Companies that patch one medium (like web pages) but ignore videos can be exposed. The DOJ and courts are clear: if a company provides information or services via video, those videos must be accessible. Nationally, there have been cases (e.g. Netflix, Harvard) where lack of captioning cost millions.
- Vendor Overpromising: Some vendors advertise “AI compliance” as if no human needed. In reality, as accessibility experts warn, AI captions often “*do not meet WCAG success criteria due to being inaccurate, incomplete...*”. Companies risk making false assumptions if they don’t audit vendor claims.
- Outdated Guidelines: Misunderstanding WCAG criteria can lead to mistakes. For example, confusing subtitles with captions (assuming translated subtitles are enough) will fail WCAG. Another error: posting a PDF of a transcript and thinking that suffices (it usually doesn’t unless it’s carefully integrated with the video in an accessible format).
- Platform Pitfalls: Some sites/tools have buggy implementations (e.g. captions not toggling). Not all video players allow user control over captions; using a non-accessible player is a mistake. Nationally, agencies like Section508.gov provide a list of “508-compliant players” to avoid this failure mode.
- Resource Overrun: Incomplete planning (e.g. not budgeting enough for human QA) can lead to either low-quality captions or skyrocketing costs. Missing deadlines is also a risk; regulators expect new content to be accessible by publication date, which has tripped up some organizations in other states.

By highlighting these pitfalls, decision-makers can avoid “easy ways out” (like auto-captioning without editing) and ensure that AI tools augment rather than replace necessary quality control.

## **10. COMPARISONS & ALTERNATIVES**

Below are comparisons of AI-driven methods versus traditional or alternative approaches to video accessibility:

Approach	Description	WCAG Compliant?	Pros	Cons
AI-generated captions only	Automated speech-to-text (no editing)	No	Very fast, low cost per minute.	Often inaccurate/incomplete; misses sound cues. Not WCAG-compliant on its own.
Human-edited captions (Hybrid)	AI + human review (closed captions)	Yes	Fast turn-around with high accuracy; captures sounds/speakers.	Moderate cost; requires skilled reviewers.
Manual captioning (no AI)	Human transcription & timing	Yes	Highest accuracy (99% standard); full non-speech detail.	Slowest and most expensive.
AI audio description (auto-assisted)	Machine suggests visual descriptions	Partial	Creates rough AD script quickly; scalable.	Limited understanding of context; errors likely.
Professional audio description	Human-written and voiced descriptions	Yes	Natural, expressive narration; fully WCAG.	Very costly (\$15–75/min); long production time.

Additional alternatives:

- **Transcripts Only:** Providing just a transcript file without captions. *Not sufficient for WCAG* if the video has audio, because WCAG requires synchronized captions for pre-recorded media (SC 1.2.2). It does help some users and SEO, but does not replace on-video captions.
- **Subtitles for Deaf and Hard-of-Hearing (SDH):** These are like captions. Using SDH files (with sound descriptions) instead of captions is functionally equivalent. Simply using “foreign language subtitles” is *not acceptable* for hearing-impaired audiences.

- Sign Language Overlay: In some contexts (e.g. TV broadcasting), a small window shows an interpreter. This is an alternative accommodation for the deaf. However, it is usually provided in addition to captions, not instead of them. As an *adjacent solution*, it may help some users but is expensive and not mandated by WCAG (WCAG treats it as optional).
- On-Screen Text (Burned-in captions): Hard-coded text in video editing. This meets caption content needs but fails WCAG’s “user-controllable” requirement (captions must be toggleable). It also makes localization harder. Generally not recommended when compliance is needed.

These comparisons underscore that hybrid AI-human approaches are generally the best practice: they achieve WCAG compliance with good efficiency. Table 1 (above) highlights the trade-offs. In particular, no current method beyond human oversight reliably ensures full accessibility, so AI should be viewed as an aid rather than a replacement.

## **11. FUTURE OUTLOOK (12–36 MONTHS)**

### **A) NEW JERSEY LENS**

**Policy and Preparation:** New Jersey organizations should prepare now for the April 2026 compliance deadline. By 2026, all public-facing NJ videos must meet WCAG 2.1 AA, so companies (even outside government) will likely adopt similar timelines. Procurement cycles for fiscal year 2026 will include accessibility clauses. NJ businesses should inventory their video content, prioritize high-impact materials (training, customer info), and plan remediation. It’s also probable that NJ’s Office of Information Technology will update guidance and maybe provide funding or incentives (as with other digital initiatives). Expect educational outreach (e.g. webinars, workshops in NJ cities) on WCAG compliance.

**Technology Trends:** Over the next 1–3 years, AI caption and transcription accuracy will improve. We anticipate that by 2028 many errors in current ASR engines will have been corrected (e.g. better handling of accents, medical terms). Adobe, AWS, Google, and others are continuously training models on more diverse speech data. This means by 2028, NJ companies might spend even less time on edits. Audio description AI is also likely to advance: for example, models

trained on visual storytelling could generate more nuanced descriptions. However, human review will still be needed for cultural and context correctness.

**Regulatory Evolution:** Nationally, watch for possible DOJ actions post-2026: if compliance lags in many places, DOJ might extend deadlines or issue stricter enforcement guidelines. On the federal level, there may be updates to Section 508 (potentially moving to WCAG 2.1 AA by 2030) that NJ-covered entities will follow. In NJ, given the civil-suit climate, private plaintiffs might start targeting video specifically (so far most suits cite websites but a precedent could emerge for video). Businesses should monitor case law.

**Market and Vendor Maturity:** We expect to see new entrants offering integrated AI accessibility pipelines. For instance, video hosting platforms might build-in real-time captioning and description features. NJ firms should evaluate these emerging tools early. Vendors will likely advertise improved accuracy and lower costs; decision-makers should verify claims. The trend will be toward more *automation* in workflows: e.g. cameras with built-in captioning microchips, or editing suites that auto-generate AD scripts. Staying on top of vendor roadmaps (asking for demos, proof-of-performance) will be crucial.

## **B) NATIONAL (USA) LENS**

**AI Evolution:** Nationally, “hallucination” issues in AI are being addressed, so we expect model improvements. Language models might soon assist in checking caption quality, not just generating it. For example, an AI could automatically flag a caption if it seems inconsistent with earlier speech. Generative AI tools may even propose initial AD scripts for review. By 2027, full episodes of some content might be captioned and “audio-described” by end-to-end AI, requiring only a final audit.

**Legal and Standards:** National oversight may tighten. The DOJ rule (Title II) is a milestone, but enforcement of Title III (public accommodations) could follow with new guidance or settlements referencing WCAG 2.1. If “level playing field” momentum continues, federal courts may adopt even newer WCAG (e.g. 2.2) in the coming decade. In Congress, there is discussion of an “ADA Title IV” for digital; if passed, it could set explicit standards (beyond WCAG).

Companies nationwide should keep an eye on any federal or state legislation in 2026–2028 that codifies digital accessibility requirements.

Automation vs. Workforce: There's a parallel trend in the labor market. As AI handles routine captioning, the human role will shift toward quality assurance and strategic roles (accessibility managers, compliance officers). Enterprises might retrain captioners as ADA specialists. We should also watch whether "gig" economy platforms emerge for crowdsourced caption checking, as hinted by some accessibility firms.

User Expectations: With tech-savvy users, even beyond compliance, consumers will expect accessibility as a standard feature. Industry leaders (e.g. Netflix, Disney+) already promote their accessibility options. Other industries (retail video, social media influencers) may follow suit. Over 12–36 months, accessible video could become a competitive norm rather than a niche compliance issue.

Overall, the trajectory is clear: accessible video will become faster and cheaper to produce thanks to AI, but the goals will rise (higher accuracy, more personalization, full inclusivity). Companies that build systems now will be ahead of both legal mandates and user demand. In NJ and the wider USA, preparing for this future means investing in AI-assisted processes today.

## 12. ACTIONABLE TAKEAWAYS (CHECKLIST FORMAT)

Below are practical checklists and questions for decision-makers and implementers, divided into New Jersey and national contexts.

### A) NEW JERSEY LENS

- Checklist for Video Projects:
  1. Inventory all existing NJ web videos. Identify which lack captions or descriptions.
  2. For new video content, *plan for accessibility up front*: include captioning and AD in the production budget.
  3. Choose video players that support WCAG features (toggleable captions, text size, contrast controls).
  4. Ensure any on-screen text meets contrast and legibility (minimum 4.5:1 ratio).
  5. Archive old content carefully: If truly historical/inactive, document why captions/AD aren't provided (one of NJ's exceptions). Otherwise, update it.

6. Establish a review workflow: assign a compliance lead to proof captions, possibly with a checklist (accuracy, timing, completeness).
- Vendor/Evaluation Questions:
    1. *WCAG Standards*: Ask vendors how they ensure WCAG 2.1 AA compliance for captions and AD. Can they cite methodology?
    2. *AI vs Human*: Inquire about their use of AI: Do they rely solely on auto-captions or combine with human editing? (WCAG compliance requires human-reviewed captions.)
    3. *Accuracy Metrics*: Request their caption accuracy rate (percent of words correct) and non-speech coverage. Do they measure against the 99% standard?
    4. *Turnaround & Cost*: Confirm pricing per minute and typical delivery times. What happens if content changes and requires rework?
    5. *Local Expertise*: For NJ projects, ask if they have experience with NJ-specific terminology (local names, acronyms) and New Jersey contexts (e.g. NJ Transit codes).
  - Implementation Steps:
    1. *Pilot Test*: Try an AI-captioning service on a small NJ video (like a department meeting) and measure effort to correct it.
    2. *Workflow Integration*: If using NJ state systems or LMS, integrate captioning tasks into the process. For example, configure automatic caption generation in the CMS, then flag for review.
    3. *Training*: Provide brief training to content producers on accessible video best practices (even share UW or Columbia guidelines tips like “[No burned-in text][38+L210-L218]”).
    4. *Legal Coordination*: Have legal/compliance consult on WCAG conformance liability – e.g. ensure vendor contracts include indemnity if captions fail to meet ADA.

5. Monitoring: Periodically test random NJ videos for accessibility (use a tool or manual check). Keep records of completed work (date, standards met) for audits.

## **B) NATIONAL (USA) LENS**

- Checklist for Video Projects:
  1. Ensure captions are accurate, complete, and synchronized for all prerecorded video (WCAG 1.2.2, 2.1 AA).
  2. Include captions on live-streaming or on-demand events whenever possible (WCAG 1.2.4).
  3. Provide audio descriptions for prerecorded videos where key visual information is not communicated in audio (WCAG 1.2.3).
  4. Test the final videos with actual disabled users or with assistive tech (screen readers, caption toggling) before release.
  5. Maintain style consistency: use consistent labeling of speakers and sounds across videos (per Section 508 advice).
- Vendor/Evaluation Questions:
  1. *Regulatory Knowledge*: Does the vendor stay updated on DOJ ADA and FCC rules? Do they confirm compliance with latest WCAG/ADA standards?
  2. *Technology & Quality*: What ASR engine do they use, and how do they validate its output? Do they employ standard caption QC processes (e.g. checking 99% accuracy)?
  3. *Accessibility Features*: Aside from captions/AD, do they offer transcripts, audio control, multi-language support? Can they adapt videos for different platforms (mobile, web)?
  4. *Past Performance*: Can they provide case studies or references from similar industries (especially government, healthcare, finance, manufacturing – which are prominent sectors)?

5. *Future-Proofing*: How will they handle WCAG 2.2 (coming in 2023+) or new standards? Do they have a roadmap for emerging needs (like VR/AR video accessibility)?
- Implementation Steps:
    1. Adopt Standards Document: Keep a concise internal doc summarizing WCAG criteria for media and update it as rules evolve.
    2. Leverage AI Tools: Evaluate AI captioning/transcription platforms (even free ones like YouTube Auto-Captions) to understand cost/effort trade-offs before committing to paid services.
    3. Automate Where Possible: For U.S. organizations, integrate captioning into video upload processes (e.g. through APIs). Set up alerts for new content without captions.
    4. Cross-Functional Team: Involve legal, IT, and content teams together. For example, have IT handle the technical integration of caption files, while legal verifies compliance.
    5. Budget Planning: Include accessibility line-items in media budgets. Nationally, build recurring costs for captioning (e.g. expect 0.2%–0.5% of media budget for high compliance).

By following these steps and asking the right questions, decision-makers can ensure that AI solutions enhance accessibility rather than introduce new gaps. This report serves as a foundation – the organization should continually refine processes as both technology and regulations advance.

Sources: Authoritative sources on accessibility standards and practices were used throughout this report, including W3C WCAG documentation, DOJ/ADA guidelines, industry analyses, and New Jersey government resources. Where specific numbers or claims could not be directly cited, they are described as estimates or best practices based on multiple industry sources.