

# AI Toolkit — Grounded Link Extracts (Batch 3)

Access date: January 27, 2026. Each entry includes a verifiable URL plus a short, AI-ingestible extract derived from the linked source.

## 1. Local newsrooms are building AI chatbots fast and cheap

**URL:** <https://www.niemanlab.org/2025/08/local-newsrooms-are-building-ai-chatbots-fast-and-cheap/>

**Source:** Nieman Journalism Lab • **Date:** Aug 25, 2025 • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “Using Zapier, the cost to build and run each chatbot came out to about \$40 per month.”

**Why this matters:** Describes a rapid, low-cost approach to building newsroom chatbots with no-code tooling and light customization—useful precedent for small newsroom pilots and cost modeling.

**AI-ingestible extract:** CISLM ran pilots where small newsrooms launched custom chatbots in weeks using Zapier; reported operating cost was about \$40/month per bot, with each newsroom tailoring the bot to its own customer service and archive-navigation needs.

## 2. How Newsrooms Are Using AI Chatbots to Leverage Their Own Reporting — and Build Trust

**URL:** <https://gijn.org/stories/newsrooms-using-ai-chatbots-leverage-reporting/>

**Source:** Global Investigative Journalism Network (GIJN) • **Date:** Aug 7, 2025 • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “It can churn out a thoughtful, accurate summary answer, with footnotes and links to FT stories that informed its response.”

**Why this matters:** Shows the “grounded chatbot” pattern: answers constrained to a publisher’s own archive with visible footnotes/links—exactly the interaction model you want for a grounded citation layer.

**AI-ingestible extract:** GIJN profiles newsroom chatbots designed to answer questions using only trusted internal reporting archives and vetted databases, producing summaries alongside footnotes and links back to the original stories.

## 3. JournalismAI Innovation Challenge Report 2024

**URL:** <https://www.journalismai.info/research/2024-innovation-challenge-report>

**Source:** JournalismAI (Polis, LSE) • **Date:** (Report page; date not shown on snippet) • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “This report captures the collective experiences of 35 news organisations across 22 countries...”

**Why this matters:** A program-level evidence base for how newsrooms designed and implemented AI tools; good for grounded “case study” citations and implementation lessons.

**AI-ingestible extract:** The report compiles case studies and practical lessons from 35 news organisations across 22 countries participating in the first JournalismAI Innovation Challenge, documenting design/testing/implementation of AI tools in workflows.

## 4. Sowt — JournalismAI Innovation Challenge 2024

**URL:** <https://www.journalismai.info/programmes/innovation/innovation-challenge-2024/sowt>

**Source:** JournalismAI (Polis, LSE) • **Date:** (Project page; date not shown on snippet) • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “This grantmaking programme enabled 35 news organisations... to experiment and implement solutions... using AI technologies.”

**Why this matters:** Pointer page for a specific Innovation Challenge grantee; useful as an authoritative directory node that links out to the 2024 report and other grantees.

**AI-ingestible extract:** Sowt’s grantee page sits within JournalismAI’s Innovation Challenge directory and links to the 2024 report and other grantees—useful for grounded navigation to primary program documentation.

## 5. TurboScribe: Transcribe Audio and Video to Text

**URL:** <https://turboscribe.ai/>

**Source:** TurboScribe • **Date:** (Product page; date not shown on snippet) • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “TurboScribe converts audio and video files to text in 98+ languages...”

**Why this matters:** A concrete example of an “AI transcription” vendor claim set (languages, pricing, privacy)—useful for a tool card with verifiable vendor statements.

**AI-ingestible extract:** TurboScribe describes itself as an AI transcription service offering transcription for audio/video in 98+ languages and lists pricing tiers; the site also claims encryption and user-controlled deletion in its security/privacy FAQ.

## 6. HappyScribe: AI Notetaker, Transcription, Subtitles & Translation

**URL:** <https://www.happyscribe.com/>

**Source:** HappyScribe • **Date:** (Product page; date not shown on snippet) • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “AI-Notetaker, transcription & subtitles available... in 120+ languages...”

**Why this matters:** Another transcription vendor with a distinct positioning (AI + human review); good for a grounded comparison entry with verifiable language coverage claims.

**AI-ingestible extract:** HappyScribe advertises AI and human-assisted services for transcription/subtitles/translation and claims coverage across 120+ languages, positioning human review as an add-on for production-ready quality.

## 7. DeepL Pro — Data security

**URL:** <https://www.deepl.com/en/pro-data-security>

**Source:** DeepL • **Date:** (Product/security page; date not shown on snippet) • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “Texts are never stored or used for model training without your consent.”

**Why this matters:** Useful for “privacy & data handling” grounded claims when recommending translation/writing tools in sensitive newsroom contexts.

**AI-ingestible extract:** DeepL’s security page describes GDPR alignment, SOC 2 Type II, and states that texts are not stored or used for model training without user consent, alongside options like BYOK and audit logs.

## 8. Introducing Whisper

**URL:** <https://openai.com/index/whisper/>

**Source:** OpenAI • **Date:** Sep 21, 2022 • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “Whisper is... trained on 680,000 hours... enabling transcription... and translation... into English.”

**Why this matters:** Primary source for what Whisper is and what it can do; ideal grounding reference for any Whisper-based workflow.

**AI-ingestible extract:** OpenAI introduces Whisper as an automatic speech recognition system trained on 680k hours of multilingual/multitask data, describing robustness benefits and capabilities including multilingual transcription and translation to English.

## 9. openai/whisper (GitHub repository)

**URL:** <https://github.com/openai/whisper>

**Source:** GitHub (OpenAI) • **Date:** (Repo; rolling) • **Type:** Web page/PDF (as applicable)

**Key excerpt (≤25 words):** “Whisper is a general-purpose speech recognition model... trained on a large dataset...”

**Why this matters:** Canonical implementation + license + README for reproducible deployment; excellent grounding reference for “run locally” instructions.

**AI-ingestible extract:** The official GitHub repository describes Whisper as a general-purpose model that supports multilingual speech recognition, translation, and language identification; it links to the paper, model card, and examples.