

Narrate

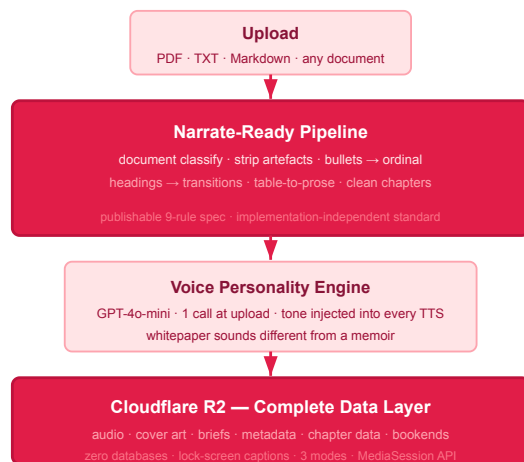
Converts any document into a narrated audiobook — voice personality per book, AI cover art, adaptive streaming, and lock-screen captions.

narrate-psi.vercel.app

THE PROBLEM

PDFs, whitepapers, and long documents sit unopened because sustained reading requires uninterrupted attention that modern life rarely provides — and audiobook platforms only serve commercially published content. Document-first sources like government notices, contracts, and circulars are completely inaccessible to anyone who cannot read due to visual impairment, low literacy, or language barriers.

ARCHITECTURE & DATA FLOW



WHAT I BUILT

- › Narrate-Ready pipeline: any uploaded file (PDF, TXT, Markdown) runs through a 9-rule preprocessing spec — document type classification, language detection, header stripping, bullet-to-ordinal conversion, table-to-prose, abbreviation expansion. Published as a canonical, implementation-independent standard.
- › Voice personality per book: one AI call at upload writes voice instructions specific to this book's genre and tone, injected into every TTS chapter call. A strategy whitepaper and a memoir sound noticeably different.
- › Adaptive buffering: a 200-chapter book starts playing in under 30 seconds. Conversion runs in the background as you listen — the same model as YouTube's progressive video loading. GlobalConversion at layout level survives all navigation.
- › AI cover art (DALL-E 3), narrated brief (full-book audio summary), "Where you are" context card (SO FAR / NEXT UP for returning listeners), and intro/outro bookends that frame every listening session.
- › Lock-screen captions in 3 modes (off / phrase clusters / full sentence) via iOS MediaSession API. No database — Cloudflare R2 is the complete data layer for metadata, audio, cover art, and chapter data.

WHAT THIS DEMONSTRATES

Product craft — architecturally complex (adaptive streaming, preprocessing pipeline, MediaSession integration) and useful enough to become a daily habit. The Narrate-Ready preprocessing spec is a publishable document-to-audio standard.

KEY NUMBERS

8+books
narrated**<30s**start, 200
chapters**3**lock-
screen
modes**0**databases
used

TECH STACK

Next.js 15

TypeScript

OpenAI TTS

GPT-4o-mini

DALL-E 3

Cloudflare R2

Vercel

MediaSession API

Sameer Bhalerao

Senior Analytics & AI
Leader · Luxembourg
11+ years · 9 years at
Amazon (L6)
Built solo · personal daily
driver

sameerbhalerao.com
linkedin.com/in/sameervb