

# LLM Scan

PUBLIC AI VISIBILITY REPORT

# www.bytedance.com

Scanned Jun 23, 2026, 08:32 UTC

OVERALL SCORE

# 50 /100

Needs Work

## Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include crawler policy and sitemap, while plain-text page access and structured data need attention. Recommended next step: add content negotiation for Accept: text/markdown on the homepage and return a markdown representation with Content-Type: text/markdown. Keep the HTML response for regular browser requests.

## Recommended next step

1. Add content negotiation for Accept: text/markdown on the homepage and return a markdown representation with Content-Type: text/markdown. Keep the HTML response for regular browser requests.
2. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.
3. Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.

## Signal breakdown

### Crawlability

Warn 10/20

Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.

### Robots.txt

Pass 15/15

robots.txt allows crawler access and includes Sitemap references.

### llms.txt

Warn 7.5/15

Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

## Sitemap

Pass 10/10

The sitemap.xml file is valid and contains URL entries.

## Markdown support

Fail 0/15

Add content negotiation for Accept: text/markdown on the homepage and return a markdown representation with Content-Type: text/markdown. Keep the HTML response for regular browser requests.

## Semantic HTML

Warn 7.1/10

Shorten the meta description to 160 characters or fewer. Add missing semantic elements: main, article, nav, footer.

## Structured data

Fail 0/10

Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.

## Content signals

Fail 0/5

Add the standard directive 'Content-Signal: ai-train=no, search=yes, ai-input=yes' to robots.txt, HTML metadata, or HTTP headers so AI systems can discover content usage preferences.

## Suggested fixes

### Fix Structured data

#### HTML

```
<script type="application/ld+json">
{
"@context": "https://schema.org",
"@graph": [
{
"@type": "Organization",
"@id": "https://www.bytedance.com/#organization",
"name": "ByteDance - Inspire Creativity, Enrich Life",
"description": "ByteDance is a technology company operating a range of
content platforms that inform, educate, entertain and inspire people across
languages, cultures and geographies.",
"url": "https://www.bytedance.com/",
"logo":
"
Boo35HAAACFlBMVEUAAACp3f8yW7U8jf8zW7U/3vJA5P8AydMzW7VP1P955908jp8AydM8jf8yWr
V5590yWrV45908jf8yW7UyW7UyW7V55909jf88jf8yW7UzW7V56N49jf80XLY+j/80XbY/lf8zZr
51lf9GXbkA2+RCY70zZsx459x5590AydI8jp8AydN4590AydM8jf8AydM8jf88jf94590AyNM8jf
8AydN5590AydM8jf8AyNIAytR5594zW7V56N09jf955948jf8AydMAytN65948jf956N56594Ayt
MAy9Q9jv976d8AytYAy9Z66N566uE/jv8AztUAy9cAytV56OSD798A0NYAydM8jp95590yW7R45t
wAydN45909jp8yWrU8jf945t0AydIyWrV5590AydIyW7UzWrU8jf8AydIzW7UzWrV459145908jf
8yW7UAydM8jp8AydM8jf9459155909jf8zW7UAydQAytMAydN56N0AydMzXLYAydM0W7UAydQ8jv
8+jv8Ay9QzW7Y9jf965t176N80W7YAydN56N156d4yW7U0XLc+jv8+kP8Ay9VAj/81XbgAzNQ1Xb
Y5XbZAKf9+7eQ7Yrp96+FDlv+F6emI/+4AydN55t15594AydM+jf8+jf8+jv8AydQ1W7QzW7UAyd
R75942XbUyWrR86OCA8uYAZtqA8eN45twyWrQAyNI8jp8bed+KAAAArnRSTLMAAfb9+wUD8HMJ/f
r59+jbldTLt6eamY+Jg3xXTkhHOHIRDQsJBwX59vbr4+DcltXRxcC7ubSsqoJuak5CKiIN+enlyb2
1lXVFEOj09NzAnJiOdGBUPDvzx8Q/s6Ob13970zsvIyMXEwcG/vbiysbGuqqeknZSTkI2Gf3ZraG
Continued in the full scan report...
```

### Fix Content signals

#### HTML

```
<meta http-equiv="Content-Signal" content="ai-train=no, search=yes,
ai-input=yes" />
<meta name="content-signal" content="ai-train=no, search=yes, ai-input=yes"
/>
```

## Full report

<https://www.llmscan.dev/scan/FoE6I8ZxMbdubRkMISkdI>