

PR #36784 完整报告

sgl-project/sglang

[Docs] Feature GLM-5.3-Flash in the popular-models banner

合并时间: 2026-08-28 12:56

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/36784>

执行摘要

本 PR 为纯文档更新, 将 GLM-5.3-Flash 加入 docs 首页与 Cookbook 首页的热门模型轮播, 同时移除 Inkling 条目, 并修正 GLM 厂商卡片链接。变更仅涉及两个文档文件, 无代码逻辑改动, 风险极低。

功能与动机

GLM-5.3-Flash cookbook 已上线, 需要将其作为热门模型推荐展示, 同时淘汰 Inkling 以保持轮播精简。PR 描述明确提及更新轮播和厂商卡片链接。

实现拆解

- 更新 `docs/src/snippets/configs/popular-models.jsx`: 在 `popularModels` 数组新增 GLM-5.3-Flash 条目, 并删除 Inkling 条目。新增条目包含完整 hero 块, 平台数 (9) 与模型配置 `supportedHardware` 一致, 文案源自模型页面开场。
- 更新 `docs/cookbook/autoregressive/intro.mdx`: 将 GLM 卡片链接从 GLM-5.2 改为 GLM-5.3-Flash。

由于轮播消费端遍历 `models.length` 且总数保持 4, 无需组件改动。无测试文件变更。

`docs/src/snippets/configs/popular-models.jsx`

热门模型轮播数据源, 新增 GLM-5.3-Flash 条目并移除 Inkling, 是本次变更的核心。

```
// 热门模型列表 (docs 首页 + Cookbook 首页轮播)
// 新增 GLM-5.3-Flash 条目, 并移除 Inkling, 保持数组长度为 4
// 平台数量 (9) 需与模型配置 supportedHardware 一致

export const popularModels = [
  // ... 其他条目 ...
  {
    name: "GLM-5.3-Flash",
    vendor: "Z.ai",
    href: "/cookbook/autoregressive/GLM/GLM-5.3-Flash",
    logo: "/cards/logos/glm.png",
    badge: "New",
    tags: ["9 platforms", "MLA + DSA + KDA hybrid", "Multimodal"],
    hero: {
      eyebrow: "Featured model • New",
```

```
    headline: "Meet GLM-5.3-Flash on SGLang",
    blurb: "Z.ai's natively multimodal MoE model — 320B total / 18B active, ...",
    tags: ["320B / 18B active", "1M context", "Text + image + video"],
    cta: "Open the GLM-5.3-Flash cookbook",
    caption: "GLM-5.3-Flash deployment guide",
  },
},
// 移除 Inkling 条目，保持总数一致
];
```

评论区精华

无 review 评论，审核人 APPROVED。

风险与影响

风险极低。唯一需要注意文案准确性（如平台数量），但已按规则核对。影响范围为 docs 首页与 Cookbook 首页的展示及 GLM 卡片链接，对用户无负面体验。

关联脉络

与近期 PR #36804（新增 Hy4-Preview 模型页）同属文档维护主线的组成部分，未来热门模型轮播可能持续更新。