- **11/25/2025**

  - 项目架构初步搭建
  - 整合supabase
  - 注册、登录、鉴权功能完成

- **11/26/2025**

  - 主要功能初步设计
    - Hacker News story schema
    - 查询 -> 提取网页关键内容（jina） -> AI 翻译，总结流程设计
    - PRD.md
  - Hacker News 异步访问API获取best, top, new stories 逻辑完成

TODO:

- 每半个小时执行一次轮询获取内容
- 设计将表结构存入supabase
- 探索使用jina取得网页信息
- 研究AI翻译总结的高性能prompt
- 将AI生成内容并入CreateArticle
- 类存入数据库

Hacker News Chinese Implementation Plan

1. Database & Repository Layer (Priority High)

- ☐ **Schema Setup**: Create `articles` table in Supabase matching PRD Section 4.
- ☐ **Repository Implementation**: Create `app/repositories/article_repository.py`.
  - `exists(hn_id: int) -> bool` : Check if article already exists (Deduplication).
  - `create(article: Article) -> Article` : Save processed article to DB.

2. Content Extraction & AI Services (Priority High)

- ☐ **Dependencies**: Add `openai` (or preferred LLM lib) and `requests` / `httpx` to `pyproject.toml`.
- ☐ **Config**: Update `app/core/config.py` to include `JINA_API_KEY` and `OPENAI_API_KEY`.
- ☐ **Content Service**: Create `app/services/content_extractor.py` using Jina Reader API.
- ☐ **AI Service**: Create `app/services/ai_service.py`.
  - Implement translation prompt.
  - Implement summarization/analysis prompt (JSON output).

3. Pipeline Integration (Priority Medium)

- ☐ **Update HNService**:
  - Integrate `ArticleRepository` to filter out existing stories.
  - Integrate `ContentService` to fetch body text.
  - Integrate `AIService` to process content.
  - Save final result to DB.

4. Scheduling (Priority Medium)

- ☐ **Scheduler**: Implement a background task (using `apscheduler` or `asyncio` loop) in `main.py` to run the fetch pipeline every 30 minutes.