

Để xây dựng **kế hoạch nghiên cứu AI R&D** với các mục tiêu cụ thể như trên, cần tham khảo và tổng hợp tài liệu từ các nguồn chất lượng, liên quan đến **mô hình ngôn ngữ nhỏ, AI nhúng, và tối ưu hóa thuật toán**. Dưới đây là các gợi ý về nguồn tài liệu và cách tìm kiếm:

## 1. Nguồn Tìm Tài Liệu Khoa Học và Báo Cáo

### a. Các trang cơ sở dữ liệu học thuật:

- **arXiv.org** (<https://arxiv.org>):
  - Là nơi cập nhật nhanh các nghiên cứu mới nhất về AI, Machine Learning, và mô hình ngôn ngữ.
  - Tìm các từ khóa như:
    - *"Small Language Models", "Efficient Language Models", "Edge AI"*.
    - *"Optimization for Resource-Constrained Devices"*.
  - Ví dụ: Tìm các nghiên cứu liên quan đến **TinyBERT, DistilBERT, MobileBERT** hoặc **LoRA (Low-Rank Adaptation)**.
- **IEEE Xplore** (<https://ieeexplore.ieee.org>):
  - Đây là nguồn tin cậy cho các bài báo được **SCI-index**. Tập trung vào **học máy, AI nhúng và tối ưu thuật toán**.
- **Google Scholar** (<https://scholar.google.com>):
  - Dễ truy cập và tìm kiếm các bài báo liên quan.
  - Dùng các từ khóa:
    - *"Efficient Transformer", "Edge AI Models", "Algorithm Optimization for Limited Hardware"*.
- **ACM Digital Library** (<https://dl.acm.org>):
  - Tốt cho nghiên cứu về AI hệ thống, tối ưu hóa thuật toán và phần cứng AI.
- **SpringerLink và Elsevier** (ScienceDirect):
  - Các bài báo có chất lượng, đa số SCI-index. Tìm kiếm từ khóa về AI nhúng, tối ưu hiệu suất, và mô hình nhỏ.

### b. Hội nghị và tạp chí AI hàng đầu

Tìm các tài liệu từ các **hội nghị AI nổi tiếng** và tạp chí học thuật để cập nhật xu hướng mới:

- **NeurIPS** (Neural Information Processing Systems)
- **ICML** (International Conference on Machine Learning)
- **CVPR** (Computer Vision and Pattern Recognition)
- **ACL** (Association for Computational Linguistics): tập trung NLP và mô hình ngôn ngữ
- **IJCAI** (International Joint Conference on Artificial Intelligence)
- **TPAMI** (IEEE Transactions on Pattern Analysis and Machine Intelligence)

## 2. Nguồn Mở từ Cộng Đồng và Tài Liệu Chính Thức

### a. GitHub

- Trên **GitHub**, nhiều dự án liên quan đến **mô hình ngôn ngữ nhỏ** hoặc **AI nhúng** được open-source.
- Tìm kiếm các repository liên quan như:
  - TinyBERT**: [huawei-noah/TinyBERT](#).
  - DistilBERT** và **MobileBERT**: Các mô hình nhẹ của Transformer.
  - LoRA (Low-Rank Adaptation)** và **quantization** để tối ưu mô hình.
  - Frameworks như **TensorRT**, **ONNX Runtime** để tăng tốc inference trên phần cứng hạn chế.

### b. Blog và tài liệu từ các công ty lớn

Các công ty công nghệ lớn thường chia sẻ nghiên cứu và tài liệu của họ:

- Google AI Blog**: Tìm thông tin về **MobileBERT**, **DistilBERT**, và các cải tiến Transformer nhẹ.
- Meta AI Blog**: Thông tin về **LLaMA**, **quantization** và AI tối ưu.
- Microsoft AI Blog**: Các tối ưu cho **ONNX** và inference AI nhúng.
- NVIDIA Developer Blog**: Chia sẻ tài liệu về tối ưu AI trên GPU và thiết bị biên.
- Hugging Face Blog**: Hướng dẫn và case study cho các mô hình ngôn ngữ nhỏ.

---

## 3. Sách và Giáo Trình

- Efficient Deep Learning** của **Tanmay Gupta**.
  - Giới thiệu các kỹ thuật giảm độ phức tạp và tối ưu mô hình.
- TinyML: Machine Learning on Ultra-Low-Power Devices** của **Pete Warden**.
  - Tập trung vào triển khai AI nhúng trên thiết bị nhỏ và IoT.
- Optimization for Machine Learning**: Các thuật toán tối ưu cho AI.

---

## 4. Các Báo Cáo và Xu Hướng Thị Trường

- Báo cáo từ **McKinsey**, **Gartner**, hoặc **Forrester** về xu hướng AI trong năm 2025.
- State of AI Report** (<https://www.stateof.ai/>): Báo cáo tổng hợp các phát triển AI nổi bật.

---

## 5. Tài Liệu Hướng Dẫn Bằng Video

- O'Reilly Learning**: Nền tảng bạn đang dùng, có nhiều sách và video hướng dẫn chuyên sâu.
- YouTube**:
  - Tìm kiếm các kênh về AI như **DeepLearning.AI**, **Two Minute Papers**, và **NVIDIA AI**.
  - Các từ khóa như: "*Edge AI Deployment*", "*Tiny AI Models*".

---

## 6. Tiếp Cận Cộng Đồng và Chuyên Gia

- Tham gia các hội nhóm trên **Reddit** (*r/MachineLearning*, *r/ArtificialIntelligence*).
- Kết nối trên **LinkedIn** để trao đổi với các chuyên gia AI.
- Tham gia các diễn đàn như **Kaggle**, **Stack Overflow** hoặc các cuộc thi AI để học hỏi thêm.

---

## Lời khuyên

- **Lập danh sách từ khóa cụ thể** như: "*mô hình ngôn ngữ nhỏ*", "*tối ưu Transformer*", "*AI nhúng*".
- **Sử dụng các công cụ quản lý tài liệu** như **Zotero** hoặc **Mendeley** để sắp xếp tài liệu tham khảo.
- **Tìm tài liệu SCI-index trước** để đảm bảo chất lượng cao cho mục tiêu của bạn.