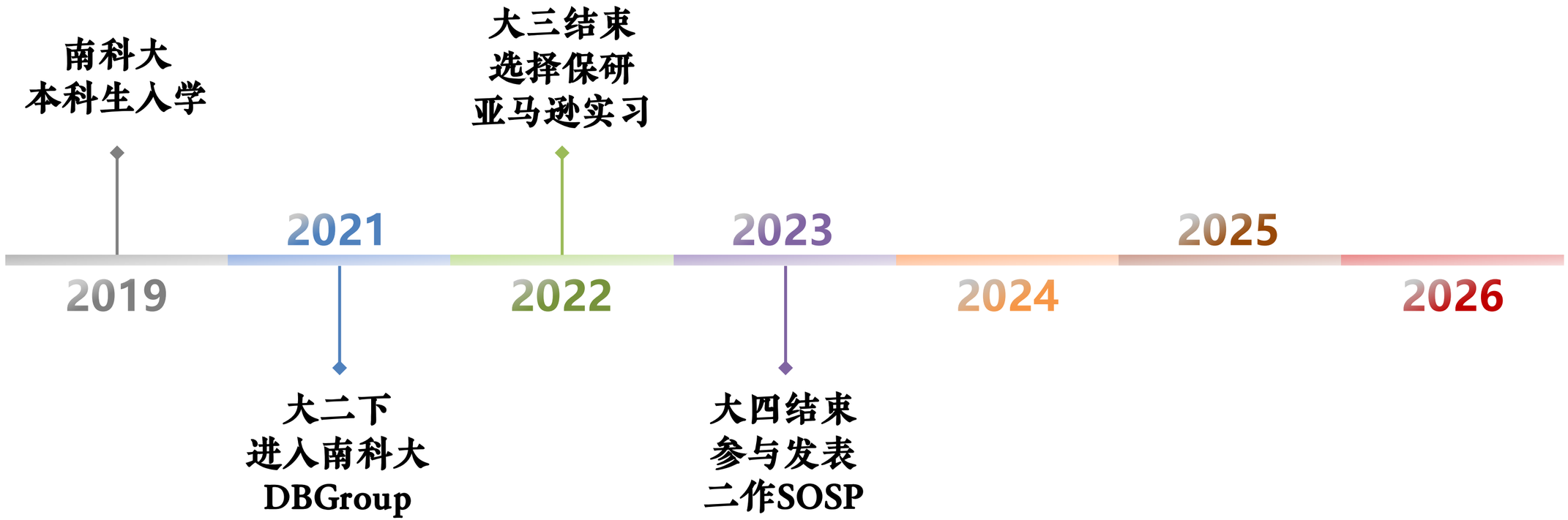
本科


南科大
本科生入学
2019
2021
大二下
进入南科大
DBGroup
大三结束
选择保研
亚马逊实习
2022
2023
大四结束
参与发表
二作SOSP
2024
2025
2026

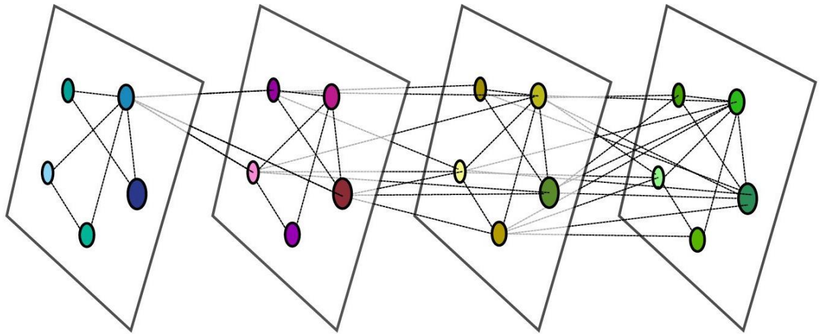
关于如何入门科研

- 科研不是课程，没有“考试”，也很难通过“刷题”提升（效率低）
- 科研并不需要智商高，而科研能力基本需要后天培养
- 主要的科研能力包括：理解力、思维深度、学习能力

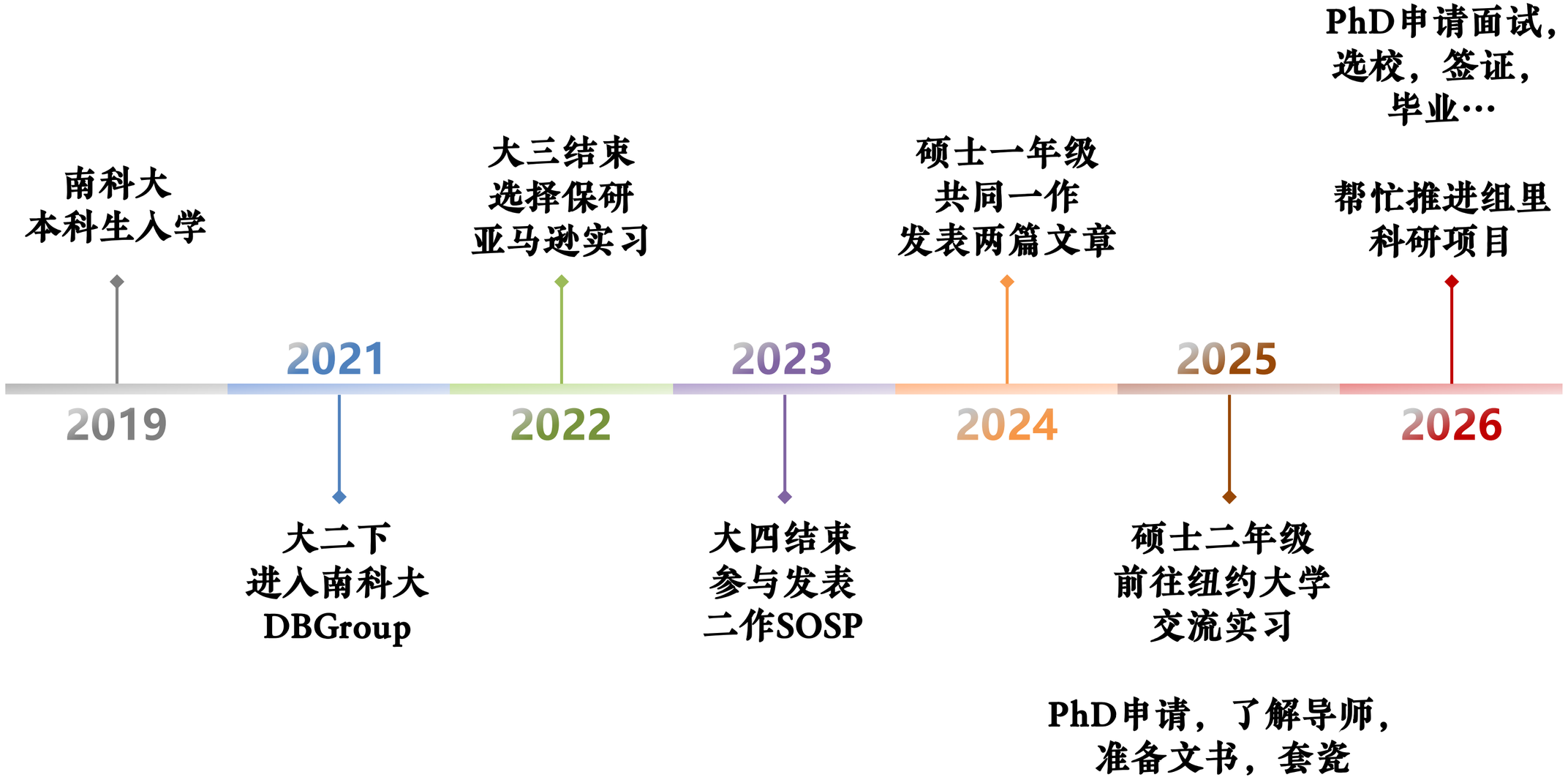
- 所以，科研入门最好的方式是先跟懂科研的人学
- 本质是跳过试错（踩坑）的过程，直接吸收凝练好的经验
- 有了基础的科研能力后，可以培养自己独特的科研风格

我在亚马逊的第一个项目是做图神经网络（GNN）GPU采样算法的加速，跟着两个博士生一年经历了从立项到投稿的完整过程，基本了解了这个领域科研的流程

在那之前，我也曾独立做过一个项目，但最终没有能够有文章产出



研究生


南科大
本科生入学
2019
2021
大二下
进入南科大
DBGroup
大三结束
选择保研
亚马逊实习
2022
2023
大四结束
参与发表
二作SOSP
硕士一年级
共同一作
发表两篇文章
2024
2025
硕士二年级
前往纽约大学
交流实习
PhD申请，了解导师，
准备文书，套瓷
PhD申请面试，
选校，签证，
毕业…
帮忙推进组里
科研项目
2026

如何独立完成科研项目

选题

了解一个领域有哪些问题/场景被做了（做到什么状态了？），
那些还没被做（还需要做什么？），可以列一个表跟导师讨论

↓

立题

当前选题是什么？出发点是什么？现有的工作都做到什么程度了？
他们还有什么问题需要解决？预期的效果是怎么样的？

↓

实现

快速实现（一两周）一个简易版本/先做粗略的数值模拟，估测
出预期能达到多少的收益（一般是一个上界），然后再完整实现

↓

验证

设计完整的主实验，消融实验，敏感性实验，首先确保在主实验
上设计的方法/系统有效果，然后再验证其他模块实验

↓

发表

将上面四步用形式化的语言描述，introduction，background，
motivation，method，evolution（让不了解领域的人也能看懂）

图神经网络训练

	是否有工作	I/O量	I/O速度	准确率
SSD场景				
	是否有工作	通信方式	跨GPU/跨机器	通信开销
分布式场景				

如何申请PhD

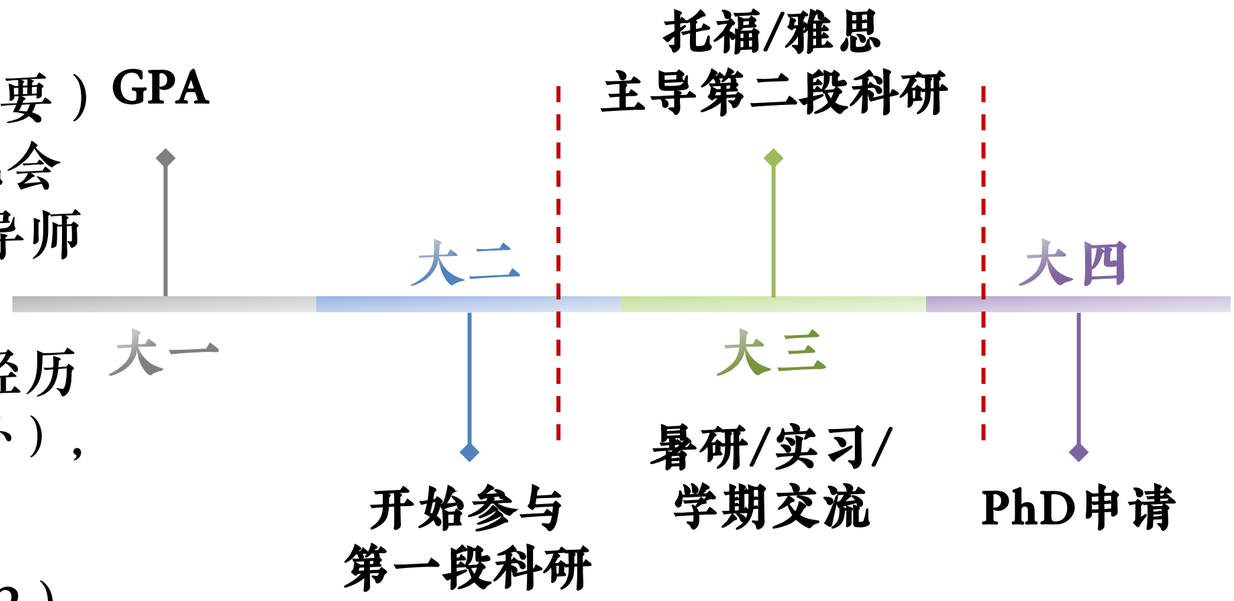
PhD ≈ 1~2篇文章 + 推荐信 + 面试

申请过程 ≈ 填报系统 + 提交推荐信 + 套瓷

重要性：推荐信 > 面试 > 文章 > 其他

- 文章：一篇主导（一作）+ 其他参与文章（可选）
- 语言：托福/雅思（两年有效），GRE（基本不需要）
- 推荐信：3封，主要来自于项目合作导师（内容也会是关于科研表现），本科生可以有一封是来自课程导师
- 套瓷：给感兴趣的导师发邮件，但很少会被回复
- 暑研/实习/学期交流：核心目标是积累一段科研经历（因此开发岗位的实习意义不太大，AI核心系统开发岗除外），有一篇可以讲的文章（不一定发表），和推荐信

一般来说会有三轮筛选：1）简历（文章，经历）；2）推荐信（打分）；3）面试（表达和思维能力）



如何申请PhD

关于收集导师信息

- 一个学校一般需要有三个以上相关/感兴趣的导师，如果太少的话，那个学校可以考虑不申请
- 列表统计信息：姓名，研究方向，（副）教授？，主页，是否套瓷，是否面试
- 列表方便管理选校/导师优先级，和更新面试/套瓷信息，知道下一步该做什么
- 把选校和导师列表发给写推荐信的老师看，这样如果有一些意向导师是认识的，老师可以写出更有针对性的推荐信

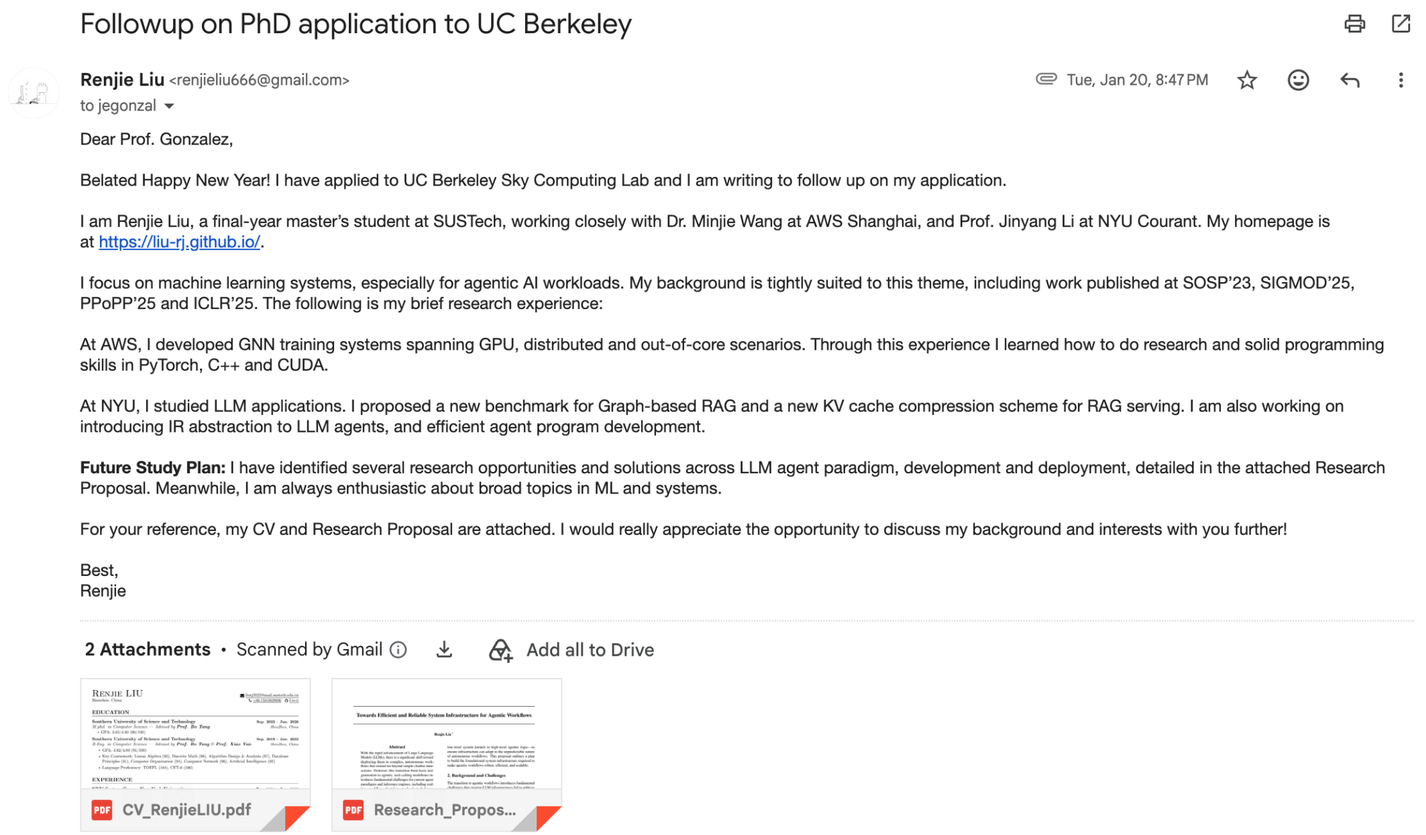
School	Faculty	Area	Title	Homepage
MIT	Tim Kraska	mlsys	Associate Professor	https://people.csail.mit.edu/kraska/
MIT	~~Song Han~~	~~mlsys & ml~~	Associate Professor	~~https://hanlab.mit.edu/~~
MIT	Adam Belay	cloud comp. & mlsys	Associate Professor	http://www.abelay.me/
MIT	Omar Khattab	nlp & ai sys	Assistant Professor	https://omarkhattab.com/
MIT	sam madden	db	Professor	https://db.csail.mit.edu/madden/
UC Berkeley	Ion Stoica	mlsys	Professor	https://people.eecs.berkeley.edu/~istoica/
UC Berkeley	Joseph Gonzalez	ml & sys	Associate Professor	https://people.eecs.berkeley.edu/~jegonzal/
UC Berkeley	Matei Zaharia	mlsys	Associate Professor	https://people.eecs.berkeley.edu/~matei/
UC Berkeley	Sewon Min	LLM & RAG	Assistant Professor	https://www.sewonmin.com/
UC Berkeley	Rishabh Iyer	systems	Assistant Professor	https://rishabh246.github.io/
UC Berkeley	Aditya Parameswaran	DB	Associate Professor	https://people.eecs.berkeley.edu/~adityagp/
CMU	Beidi Chen (ECE)	ml & mlsys	Assistant Professor	https://www.andrew.cmu.edu/user/beidic/
CMU	Zhihao Jia	mlsys	Assistant Professor	https://www.cs.cmu.edu/~zhihaoj2/
CMU	Tianqi Chen	mlsys	Assistant Professor	https://tqchen.com/
CMU	Rashmi Vinayak	distributed system	Associate Professor	https://www.cs.cmu.edu/~rvinayak/#
CMU	Tim Dettmers	ml & mlsys	Assistant Professor	https://timdettmers.com/
CMU	Phillip B. Gibbons	llm & sys	Professor	https://www.cs.cmu.edu/~gibbons/
CMU	Eric Xing	mlsys	Professor	https://www.cs.cmu.edu/~epxing/
Stanford	Azalia Mirhoseini	ml & sys	Assistant Professor	http://www.azaliamirhoseini.com/
Stanford	Christos Kozyrakis	mlsys	Professor	https://web.stanford.edu/~kozyraki/
Stanford	Christopher Re	mlsys	Professor	https://cs.stanford.edu/~chrismre/
Stanford	Jure Leskovec	graph	Professor	https://cs.stanford.edu/people/jure/
UW	Arvind Krishnamurthy	mlsys	Professor	https://sites.google.com/cs.washington.edu/arvind/
UW	Baris Kasikci	mlsys	Associate Professor	https://homes.cs.washington.edu/~baris/
UW	Stephanie Wang	mlsys	Assistant Professor	https://stephanie-wang.github.io/
UW	Luis Ceze	mlsys	Professor	https://homes.cs.washington.edu/~luisceze/
ETHz	Ana Klimović	mlsys	Assistant Professor	https://anakli.inf.ethz.ch/
ETHz	~~Torsten Hoefler~~	~~super computing~~	Professor	~~https://htor.inf.ethz.ch/~~
NYU	Jinyang Li	mlsys	Professor	https://jinyangli.github.io/
NYU	Saiqian Zhang	mlsys	Assistant Professor	https://www.saiqianzhang.com/
Cornell	Udit Gupta	AI systems	Assistant Professor	https://ugupta.com/index.html
Cornell	Rachee Singh	network	Assistant Professor	https://www.racheesingh.com/
Cornell	Mohamed S. Abdelfattah	ml & sys	Assistant Professor	https://www.mohsaied.com/
Cornell	Jiaxin Lin	network	Assistant Professor	https://www.csl.cornell.edu/~jiaxinl/
Gatech	Anand Iyer	network & sys (P3)	Assistant Professor	https://www.anand-iyer.com/
Gatech	Francisco Romero	video & mlsys	Assistant Professor	https://www.franciscoromero.com/
Gatech	Yingyan (Celine) Lin	ml & accelerator	Associate Professor	https://eiclab.scs.gatech.edu/pages/team.html#PI
Gatech	Kexing Rong	db	Assistant Professor	https://kexinrong.github.io/
Gatech	Alexey Tumanov	mlsys	Associate Professor	https://faculty.cc.gatech.edu/~atumanov/
Princeton	Ravi Netravali	network & mlsys	Associate Professor	https://www.cs.princeton.edu/~ravian/
Princeton	Tri Dao	ml & sys	Assistant Professor	https://tridao.me/
Princeton	Karthik R. Narasimhan	llm & agent	Associate Professor	https://karthikncode.github.io/
Princeton	Jialin Ding	db	Assistant Professor	https://jialinding.github.io/
UPenn	Vincent Liu	network & mlsys	Associate Professor	https://vincen.tl/
UPenn	Ryan Marcus	db	Assistant Professor	https://rmarcus.info/blog/
UPenn	Josh Fried	os for datacenter	Assistant Professor	https://joshfried.io/
UIUC	Minjia Zhang	mlsys	Assistant Professor	https://minjiazhang.github.io/
UIUC	Fan Lai	mlsys	Assistant Professor	https://www.fanlai.me/
UIUC	Jiaxuan You	mlsys & GNN	Assistant Professor	https://cs.stanford.edu/people/jiaxuan/
UIUC	Lingming Zhang	software+ml	Associate Professor	https://lingming.cs.illinois.edu/
Rice	Yuke Wang	mlsys	Assistant Professor	https://www.wang-yuke.com/
Rice	Jiarong Xing	mlsys & agent	Assistant Professor	https://jxing.me/
Rice	Eugene Ng	mlsys	Professor	https://www.cs.rice.edu/~eugeneng/
U of Edinburgh	Luo Mai	mlsys	Assistant Professor	https://luomai.github.io/
Columbia	Engene Wu	db	Associate Professor	https://www.eugenewu.net/
Columbia	Kostis Kaffes	ai system	Assistant Professor	https://www.cs.columbia.edu/~kkaffes/
Columbia	Junfeng Yang	security	Professor	https://www.cs.columbia.edu/~junfeng/
Columbia	Zhou Yu	ai agent	Associate Professor	https://www.cs.columbia.edu/~zhouyu/
Columbia	ASAF CIDON	db & os	Associate Professor	https://www.asafcidon.com/
UCSD	Yufei Ding	mlsys	Associate Professor	https://yufeiding.ucsd.edu/people/prof-yufei-ding
UCSD	Hao Zhang	mlsys	Assistant Professor	https://cseweb.ucsd.edu/~haozhang/
UCSD	Dan Fu	mlsys	Assistant Professor	https://danfu.org/
UCSD	Yiying Zhang	mlsys & agent	Associate Professor	https://cseweb.ucsd.edu/~yiying/
UCSD	Zhijian Liu	mlsys	Assistant Professor	https://zhijianliu.com/
UCSD	Jishen Zhao	mlsys & arch	Professor	https://cseweb.ucsd.edu/~jzhao/index.html

如何申请PhD

套瓷信应该包含什么内容?

- 首先说明自己提交了申请，来信是followup询问能不能聊一下
- 一段话介绍自己，学校，经历，主页
- 自己的研究方向，发表的文章
- 展开讲跟老师相关的经历，成果和收获
- 简单概述PhD想做什么，为什么跟老师和课题组相关
- 附上简历，和其他没在申请系统里提交的材料（如有）

总体来说，套瓷的回复率较低（只有三四个老师回复了我），但能让老师注意到你在系统上的申请材料，可能会给你发面试



Followup on PhD application to UC Berkeley

Renjie Liu <renjieliu666@gmail.com>
to jegonzal

Tue, Jan 20, 8:47 PM

Dear Prof. Gonzalez,

Belated Happy New Year! I have applied to UC Berkeley Sky Computing Lab and I am writing to follow up on my application.

I am Renjie Liu, a final-year master's student at SUSTech, working closely with Dr. Minjie Wang at AWS Shanghai, and Prof. Jinyang Li at NYU Courant. My homepage is at https://liu-rj.github.io/.

I focus on machine learning systems, especially for agentic AI workloads. My background is tightly suited to this theme, including work published at SOSP'23, SIGMOD'25, PPoPP'25 and ICLR'25. The following is my brief research experience:

At AWS, I developed GNN training systems spanning GPU, distributed and out-of-core scenarios. Through this experience I learned how to do research and solid programming skills in PyTorch, C++ and CUDA.

At NYU, I studied LLM applications. I proposed a new benchmark for Graph-based RAG and a new KV cache compression scheme for RAG serving. I am also working on introducing IR abstraction to LLM agents, and efficient agent program development.

Future Study Plan: I have identified several research opportunities and solutions across LLM agent paradigm, development and deployment, detailed in the attached Research Proposal. Meanwhile, I am always enthusiastic about broad topics in ML and systems.

For your reference, my CV and Research Proposal are attached. I would really appreciate the opportunity to discuss my background and interests with you further!

Best,
Renjie

2 Attachments • Scanned by Gmail Add all to Drive

CV_RenjieLIU.pdf

Research_Propos...

如何准备面试（申请文书）

SOP

Statement of Purpose

相当于套瓷信+简历的扩充

其中有两段是重点：

1. Research Vision（研究视野，一般出现在开头）

PS

Personal Statement

个人陈述

The rapid growth of deep learning has fueled the success of transformative AI applications across domains. However, new ML algorithms are typically developed with little consideration for cost-effectiveness, making them impractical to be directly deployed in real production environments. Many attempts have been proposed to solve this, among which, three directions excite me the most: 1) develop new abstractions for emerging algorithmic patterns to enable holistic optimization opportunities; 2) apply fine-grained system techniques tailored to the algorithm and hardware characteristics; and 3) jointly consider different components in end-to-end ML pipelines for better resource utilization and scalability. Over the past few years, I have done research in several projects spanning these themes. For my Ph.D. study, I want to take a broader perspective, while still solving system challenges related to the above three aspects.

Slides

面试使用的PPT

项目整理

如何准备面试（申请文书）

SOP

Statement of Purpose

相当于套瓷信+简历的扩充

其中有两段是重点：

1. Research Vision（研究视野，一般出现在开头
2. Research Plan（研究计划，一般出现在结尾）

PS

Personal Statement

个人陈述

Research Plan: In my future PhD study, I am eager to continue research on system and algorithm optimizations for AI pipelines, particularly for LLM agents. I believe that **versatile, efficient, and robust agent systems** are key to unlocking future AI developments. Specifically, 1) current LLM agents are intrinsically unstable, and thus new system abstractions for better agent program auto-tuning and reliable online execution are required; 2) agent programs are observed to have a rigorous development phase, where the agent's capabilities are refined by interleaved evaluation and optimization, and thus this stage necessitates efficient and robust system support; 3) the deployment of agentic workflows consists of both LLM generation and external function calls, and thus joint optimizations on end-to-end resource utilization and scheduling are needed. This integrated approach—designing better abstractions and optimizing their execution workflows—is the line of research that I am most passionate about and wish to pursue in my PhD. In general, I am always excited to explore topics broadly in building cutting-edge AI infrastructure and applications.

Slides

面试使用的PPT

项目整理

如何准备面试（申请文书）

SOP

Statement of Purpose

相当于套瓷信+简历的扩充

其中有两段是重点：

1. Research Vision（研究视野，一般出现在开头）
2. Research Plan（研究计划，一般出现在结尾）

https://cs-sop.notion.site/

PS

Personal Statement

综合个人陈述

讲自己的成长经历，怎么开始科研，遇到了哪些挫折，如何解决，为什么决定读PhD

突出面对挫折，教学经历

Slides

面试使用的PPT

科研项目汇总

其实是梳理面试思路的过程，怎么样把每一段项目经历讲述清楚

重点是未来想做什么，可以占比较大的篇幅讲

实际上我大部分面试都没有用slides…

如何选校

- 城市：安全性，实习、就业机会
- 课题组氛围：导师和科研风格，多和组里高年级学生聊一聊
- 毕业要求：是否是有严格毕业标准，还是看导师的主观感觉？组里学生平均毕业年限？
- 导师是否创业：对科研项目的投入时间…

选校的核心：不是挑选出最好的选择，而是筛选掉不适合自己的选择

Image from https://www.zhihu.com/tardis/sogou/qus/40040178

选校之后的事情

- 留学生身份：SEVIS ID，I-20，F-1签证（D/S，CPT，OPT政策变化）
- 签证：尽早约面签时间，一般处理时间两个月
- 选导师：一些学校的offer可能不指定导师，可以暑假开始着手跟意向导师聊一聊

- 租房，注册，选课，RA…