PROJECT IDEAS · PHD

PhD topics in computer *science*

Good PhD topics in computer science start from a question that recent surveys still call open, that you can test on public data with the compute you will have, and that your supervisor can guide.

Last updated 2026-10-11

Below are **16 topic ideas in five research areas**: trustworthy and efficient AI, LLMs and Indian-language NLP, medical imaging, networks and security, and data systems. Each is written as a research question with why it is open, a public dataset or benchmark and a feasibility note, and each area lists published surveys with checked DOIs. They are ideas to discuss with a supervisor, not approved topics.

KEY POINTS

- **16 topic ideas in five research areas:** trustworthy and efficient AI, LLMs and Indian-language NLP, medical imaging, networks and security, and data systems.
- **Each idea is a research question** with why it is open, a public dataset or benchmark and a feasibility note.
- **A topic has to last three to six years.** The UGC's 2022 PhD regulations set a minimum of three years including coursework and a maximum of six before any extension.
- **Five checks before you commit:** novelty, data access, compute, supervisor fit and publishable steps.
- **These are ideas to discuss with a supervisor,** not approved topics. The survey DOIs listed with each area were checked against Crossref on 2026-09-29.

How do you judge a PhD topic before you commit?

A PhD topic has to survive three to six years: the UGC's 2022 PhD regulations set a minimum of three years including coursework and a maximum of six before any extension. Test every candidate against these five checks before you take it to a supervisor.

Five checks for a PhD topic

CHECK	QUESTION TO ASK	WARNING SIGN
Novelty	Has a recent survey or paper already closed this gap?	Several papers from the last three years share your working title
Data access	Can you get the data this year, legally and in full?	The data is available only "on request" and nobody has replied
Compute	Will the experiments run on the hardware you will actually have?	The baseline alone needs weeks on many GPUs
Supervisor fit	Has your supervisor published near this area?	No one in the department can review your method
Publishable steps	Can it give two or three results you can publish along the way?	Everything depends on one result that arrives at the end

Duration from the UGC PhD Regulations, 2022. Your university's ordinance may add its own rules.

Topic ideas, not approved topics or delivered work

Every question below is an idea to discuss with your supervisor and check against the newest literature. None is a registered topic, and none comes from PhD work we have delivered. Our case studies are M.Tech and M.E. projects. Survey DOIs were checked against Crossref on 2026-09-29.

What are open PhD topics in trustworthy and efficient AI?

Many PhD topics in artificial intelligence and machine learning now sit here: not a new model, but a model that must also be explainable, small or private. Knowing when an input lies outside the classes a model was trained on belongs here too; our unknown class detection guide covers the out-of-distribution scores and metrics at project scale.

Topic ideas: trustworthy and efficient AI

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY
TOPIC IDEA **Do saliency maps show what an image model really uses, and can faithfulness be trained in rather than checked afterwards?**	Explanation methods often disagree, and so do the metrics that score them; how to evaluate explanations is still unsettled.	ImageNet or CheXpert, scored with the Quantus toolkit's faithfulness metrics	One GPU is enough; the hard part is the evaluation design.
TOPIC IDEA **How far can a vision transformer be quantised and pruned for an edge device, and can the accuracy loss be predicted layer by layer?**	Results at 4 bits and below vary widely between models, and measured speed-ups on real hardware often differ from the theoretical ones.	ImageNet-1k or a subset, with latency measured on a Raspberry Pi 5 or Jetson-class board	Needs GPU time for quantisation-aware training, plus the target board.
TOPIC IDEA **How can hospitals train one model together when their data differ, without the smallest site losing accuracy?**	Kairouz et al. list heterogeneous client data, fairness across clients and privacy guarantees among federated learning's open problems.	FLamby, a cross-silo healthcare benchmark with natural splits by hospital, region or scanner	Simulated on one machine with FLamby's own strategies, or Flower with some glue code; each FLamby dataset has its own licence to accept.

Where to read first

1. S. Ali *et al.*, "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence," *Information Fusion*, vol. 99, 2023, Art. no. 101805. doi:10.1016/j.inffus.2023.101805
2. P. Kairouz *et al.*, "Advances and Open Problems in Federated Learning," *Foundations and Trends in Machine Learning*, vol. 14, pp. 1–210, 2021.

doi:10.1561/2200000083

3. A. Gholami *et al.*, "A Survey of Quantization Methods for Efficient Neural Network Inference," *Low-Power Computer Vision*, CRC Press, pp. 291–326, 2022. doi:10.1201/9781003162810-13

Which PhD research topics in LLMs and Indian-language NLP are open?

Large language models change quickly, so pick a question that stays useful when the next model arrives: grounding, evaluation and Indian languages all qualify.

Topic ideas: LLMs and Indian-language NLP

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY
TOPIC IDEA **Can retrieval-augmented generation give grounded answers in Hindi, Marathi or Tamil when the documents and the question are in different languages?**	Most RAG evaluation is in English; cross-lingual retrieval and faithfulness scoring for Indian languages are far less studied.	IndicXTREME (AI4Bharat) IndicQA question-answering and FLORES sentence-retrieval tasks, plus your own document set	Open 7–8B models run on one 24 GB GPU with 4-bit weights; building a clean evaluation set is the slow part.
TOPIC IDEA **How can hallucinations be measured and reduced in one high-stakes domain, such as Indian legal or medical text, without a bigger model?**	Hallucination surveys review many detection and mitigation methods but still list open challenges, and few test sets target Indian legal or medical text.	TruthfulQA for general checks, plus a domain set you build and annotate	Annotation effort, not compute, is the bottleneck.
TOPIC IDEA **Do tokenisers built mainly for English make Indian-language text longer and costlier, and does extending the vocabulary fix accuracy as well as cost?**	Token counts differ sharply between scripts, and the effect of vocabulary extension on downstream accuracy is still being measured.	FLORES-200 parallel sentences for token counts; IndicXTREME for accuracy	Token analysis runs on a laptop; continued pre-training needs several GPUs.

Where to read first

1. L. Huang *et al.*, "A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions," *ACM Transactions on Information Systems*, vol. 43, 2025. doi:10.1145/3703155
2. L. Qin *et al.*, "A survey of multilingual large language models," *Patterns*, vol. 6, 2025, Art. no. 101118. doi:10.1016/j.patter.2024.101118
3. Y. Huang and J. Huang, "A Survey on Retrieval-Augmented Text Generation for Large Language Models," *ACM Computing Surveys*, vol. 58, 2026. doi:10.1145/3805774

IndicXTREME is described in S. Doddapaneni *et al.*, ACL 2023, doi:10.18653/v1/2023.acl-long.693.

Shortlisted a topic?

Share your area, your supervisor's field and your timeline. The consultation is free.

Message before 7 pm for a reply today.

Free consultation → WhatsApp

What are good PhD topics in medical imaging?

Medical imaging has public data and clear clinical stakes, but access rules and ethics approval shape what a three-year PhD can do.

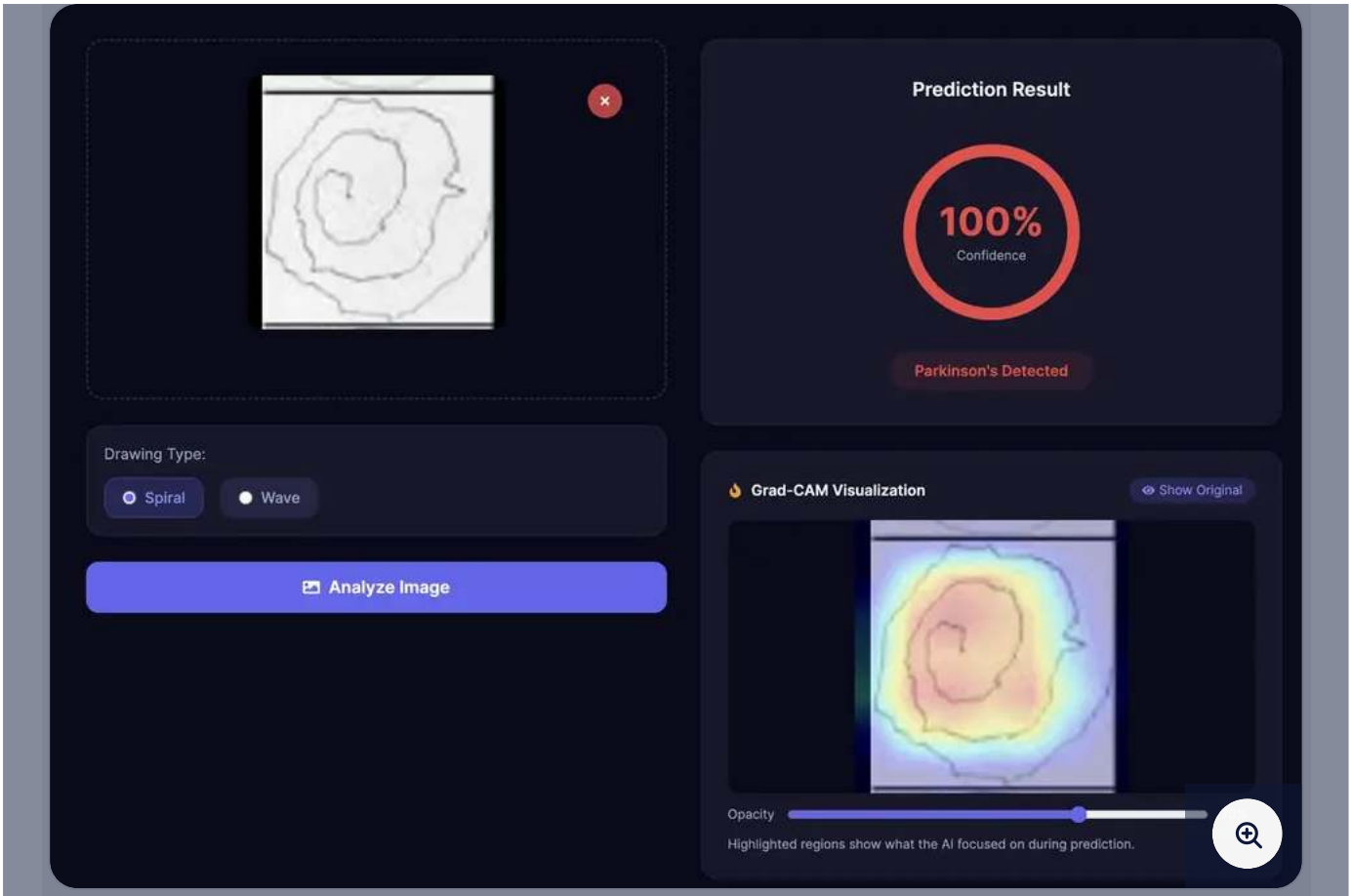
Topic ideas: medical imaging

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY
TOPIC IDEA **How many labels does self-supervised pre-training on unlabelled scans save, and does the saving hold across modalities?**	Reported gains vary with modality, dataset size and evaluation protocol, so there is no reliable rule for when it pays off.	MedMNIST v2 for fast small-scale tests; CheXpert for chest X-rays	Pre-training needs days of GPU time; MedMNIST keeps early experiments cheap.
TOPIC IDEA **Why do classifiers trained at one hospital fail at another, and can the drop be predicted before deployment?**	Careful benchmarks such as DomainBed have found that many domain-generalisation methods do not clearly beat standard training.	Train on CheXpert, test on MIMIC-CXR	One or two GPUs; MIMIC-CXR needs PhysioNet credentialing and a data-use agreement.
TOPIC IDEA **Can a model tell when to refer a case to a clinician, and do its explanations help or mislead the reader?**	Calibrated uncertainty and useful explanations are rarely evaluated together, and even more rarely with clinicians.	HAM10000 dermoscopy images (Harvard Dataverse)	One GPU; a reader study with clinicians needs ethics approval.

Where to read first

1. F. Shamshad *et al.*, "Transformers in medical imaging: A survey," *Medical Image Analysis*, vol. 88, 2023, Art. no. 102802. doi:10.1016/j.media.2023.102802
2. R. Krishnan, P. Rajpurkar and E. J. Topol, "Self-supervised learning in medicine and healthcare," *Nature Biomedical Engineering*, vol. 6, pp. 1346–1352, 2022. doi:10.1038/s41551-022-00914-1
3. K. Zhou *et al.*, "Domain Generalization: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, pp. 4396–4415, 2023. doi:10.1109/TPAMI.2022.3195549



Web app A Grad-CAM heatmap from a delivered M.Tech project, Parkinson's screening from drawings, not PhD research. Whether heatmaps like this help or mislead a reader is the third question above.

Which network, IoT and security topics are worth a PhD?

Security topics age well because attackers keep moving. The strongest questions test a method on data it was not tuned for.

Topic ideas: networks, IoT and security

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY
TOPIC IDEA **Can modelling devices and flows as a graph catch IoT attacks that per-flow models miss, at a cost a gateway can afford?**	Graph-based detectors are young, and their cost on constrained hardware is rarely reported.	CICIoT2023 and TON_IoT	A CPU or one GPU; building graphs from raw traffic is the main work.

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY
TOPIC IDEA **Do deepfake detectors still work on generators they never saw, and on images compressed by social media?**	Detection accuracy often drops on unseen manipulation methods and compressed media.	FaceForensics++ and Celeb-DF	Both datasets are released through a request form; training needs a GPU.
TOPIC IDEA **When should an inference task run on the device, the edge or the cloud, as latency, energy and network conditions change?**	Many schemes are tested in simulation with fixed conditions; results on real, changing networks are scarce.	Timings measured on two or three real devices, plus a simulator such as iFogSim	Low compute; needs careful measurement on real hardware.

Where to read first

1. T. Bilot *et al.*, "Graph Neural Networks for Intrusion Detection: A Survey," *IEEE Access*, vol. 11, pp. 49114–49139, 2023. doi:10.1109/ACCESS.2023.3275789
2. M. S. Rana *et al.*, "Deepfake Detection: A Systematic Literature Review," *IEEE Access*, vol. 10, pp. 25494–25513, 2022. doi:10.1109/ACCESS.2022.3154404
3. Y. Wang *et al.*, "End-Edge-Cloud Collaborative Computing for Deep Learning: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 26, pp. 2647–2683, 2024. doi:10.1109/COMST.2024.3393230

What are PhD research topics in data science and systems?

Data and systems topics suit scholars without large GPU budgets: most of the work is careful data handling and measurement, not training.

Topic ideas: data science and systems

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY
TOPIC IDEA **How can a deployed model detect drift and adapt without labels, and without forgetting what it learned?**	Drift detection and continual learning are usually studied apart; label-free adaptation on real temporal data remains weak.	Wild-Time, a benchmark of real distribution shift over time	One GPU.

RESEARCH QUESTION	WHY IT IS OPEN	DATA OR BENCHMARK	FEASIBILITY
TOPIC IDEA **Can maintenance models trained on rare failures give explanations that engineers find useful and correct?**	Failure data are highly imbalanced, and explanations are seldom checked with the engineers who act on them.	NASA C-MAPSS turbofan simulation data	A CPU is enough; a study with engineers is the hard part.
TOPIC IDEA **How can lending or hiring models be audited for fairness in India when attributes such as caste or region are not recorded?**	Most fairness methods assume the protected attribute is known, and public Indian datasets that include such attributes are rare.	Folktables (US census data) to develop methods, then a partner's data under agreement	Data access and ethics approval, not compute, decide feasibility.
TOPIC IDEA **Where do learned cardinality and cost models beat a database's hand-tuned optimiser, and why do they fail elsewhere?**	Leis *et al.* found that cardinality estimates routinely carry large errors and matter far more to plan quality than the cost model; robustness of learned estimators to data updates and new queries is still being studied.	The Join Order Benchmark on the IMDb data, in PostgreSQL	CPU and disk heavy; most work needs no GPU.

Where to read first

1. J. Lu *et al.*, "Learning under Concept Drift: A Review," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, pp. 2346–2363, 2019. doi:10.1109/TKDE.2018.2876857
2. L. Wang *et al.*, "A Comprehensive Survey of Continual Learning: Theory, Method and Application," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, pp. 5362–5383, 2024. doi:10.1109/TPAMI.2024.3367329
3. N. Mehrabi *et al.*, "A Survey on Bias and Fairness in Machine Learning," *ACM Computing Surveys*, vol. 54, pp. 1–35, 2022. doi:10.1145/3457607

The benchmark comes from V. Leis *et al.*, "How good are query optimizers, really?", *Proc. VLDB Endowment*, vol. 9, 2015, doi:10.14778/2850583.2850594. Our delivered M.E. / M.Tech predictive-maintenance project shows the master's-level version of the second question.

Working outside core computer science? Our PhD research topics in engineering give 20 more problems across AI, networks, electronics, energy and applied fields,

each with an expected output and a first step.

How do you narrow a topic into a synopsis?

A research area is not a topic yet. These steps turn one question from this page into something a committee can approve.

1. **Pick one question and one dataset**
   Choose the question you can start testing this month, on data you already have access to.
2. **Read the surveys, then build a literature matrix**
   Start with the surveys above, then put 20 to 30 recent papers in one table: method, data, result and limitation. Search Shodhganga for Indian theses on the same problem.
3. **Write the gap in two sentences**
   Name what is missing and point to the papers that show it. If you cannot, the topic is still an area.
4. **Turn the gap into three or four objectives**
   Each should be testable with a named dataset, method and metric.
5. **Write the proposal, then the synopsis**
   For admission, follow our research proposal for PhD guide; at registration, the PhD synopsis format guide covers the fuller document.

Studying at master's level first? The AI and machine learning project ideas and our M.Tech projects page cover topics sized for one or two semesters.

Frequently asked questions

What are the latest PhD topics in computer science for 2026?

Active areas include trustworthy and efficient AI, LLMs and Indian-language NLP, medical imaging, network and IoT security, and data systems. This page gives 16 topic ideas as research questions, each with a dataset and a feasibility note, to test against the newest papers with your supervisor.

How do I choose a PhD research topic in computer science?

Check five things: novelty against the last few years of papers, access to data, the compute you will have, your supervisor's expertise, and whether the topic can yield two or three publishable results. Read two or three recent surveys first; their open-problems sections are the quickest route to a real gap.

Which PhD topics in machine learning and AI are good for 2026?

Topics that pair a model with a hard constraint: explanations that must be faithful, models that must fit an edge device, training that must keep data private, or answers that must stay grounded in Indian languages. Each has public benchmarks, so progress can be measured.

Can I do a PhD in computer science without large GPUs?

Yes, if you choose the question with that limit in mind. Evaluation studies, tokeniser analysis, drift detection, graph-based intrusion detection and database optimisation can run on one GPU or a CPU. Pre-training large models cannot, so leave those questions to groups with the hardware.

How do I check whether a PhD topic has already been done in India?

Search Shodhganga, INFLIBNET's repository of Indian PhD theses, and Shodhgangotri for synopses already registered, alongside Google Scholar and IEEE Xplore for recent papers.

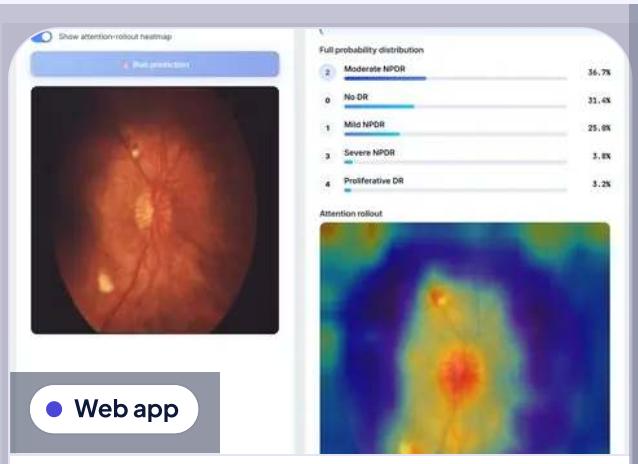
Can you help me finalise a PhD topic?

Yes. PhD guidance starts with a short list of candidate topics, each with notes on novelty, data and feasibility, for you to discuss with your supervisor. The final choice stays with you and your supervisor, and your topic is never shared.

DELIVERED WORK

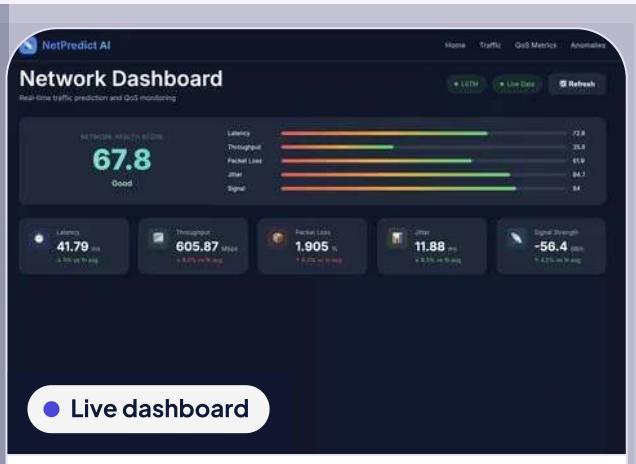
Research-style builds from delivered M.Tech work

Not PhD research: M.Tech and M.E. projects with a baseline, a measured result and a paper or report. Shown with names and institute details removed.



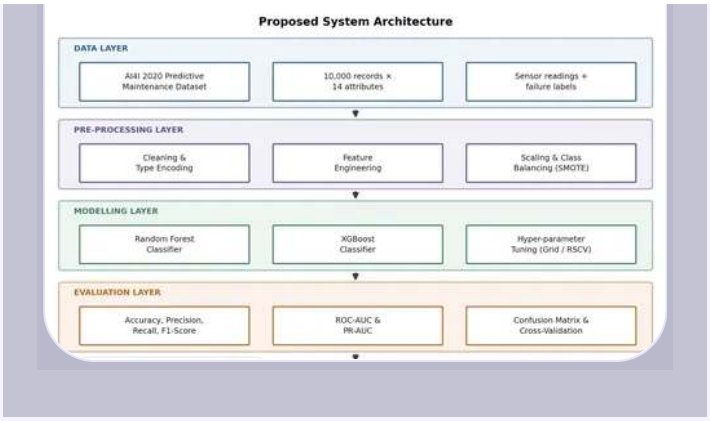
M.Tech Medical imaging · Explainable AI

Explainable diabetic retinopathy grading with a ViT ensemble



M.Tech Deep learning · Networks

Network traffic forecasting and a QoS dashboard



[See all 8 case studies →]

RELATED SERVICES

Help from topic to thesis

PhD guidance
Topic, proposal and synopsis, stage by stage. →

Research papers & publication
IEEE, Springer or journal format, with reviewer replies. →

PhD thesis writing
Chapter-by-chapter support in your format. →