From a Broad Interest to a Researchable Question

Unit I · Lecture I.3 — Narrowing, scoping,

and the five problem-statement templates you will actually use

Dr. Mohsin Furkh Dar

CSEG3060 · Research Methodology in Computer Science
School of Computer Science, UPES · B.Tech CSE 5th Semester



Three of your own statements — which one is researchable?

A Divyanshu Rawat

"Does machine learning improve fraud detection over rule-based systems under real-world banking conditions?"

B Harshit Sinha

"Does migrating to Server-Side Rendering improve First Contentful Paint?"

C Rishit Raj

"Does active learning improve model performance over random sampling, under limited labelled data?"

Hands up — exactly one

Which of these could you actually finish, and defend, in two semesters?

All three are real submissions from this batch, quoted exactly as written.

What you should be able to do by 55 minutes from now

LO1 · Narrow
Take a broad interest down to a single researchable question using a repeatable funnel.

LO2 · Calibrate
Recognise when a question is too broad *and* when it has become too narrow to matter.

LO3 · Template
Choose among five standard problem-statement forms and write yours in one of them.

LO4 · Cost
Estimate whether the experiment your statement implies fits your actual time and hardware.

Maps to CO1 *Identify and formulate research problems using defined criteria and characteristics.*

Carry-in: I.1 gave you the four ingredients (baseline · dataset · metric · constraint). I.2 gave you five sources. Today: how wide the question should be.

Hook: two projects, opposite fates

IBM Watson for Oncology vs AlphaFold

Watson for Oncology · 2013–2017

The framing. "Use AI to help cure cancer." Partnership with a leading cancer centre; enormous publicity.

The scope. Never pinned down. Which cancers? Which decision? Compared against which standard? Measured how? Across which patient population?

The outcome. A university audit reported spending of the order of **$62 million** before the project was shelved in 2017. Journalists later documented unsafe or unusable recommendations in some cases.

Nothing was ever falsifiable, so nothing was ever settled.

AlphaFold · 2016–2020

The framing could have been "use AI to solve biology". It was not.

The scope they chose. One task: predict 3D protein structure from amino-acid sequence. One benchmark: **CASP**. One metric: **GDT_TS**. One deadline: the next CASP round.

The outcome. CASP13 in 2018 — clear lead. CASP14 in 2020 — a median GDT_TS around **92**, comparable to experimental methods. Organisers called the problem substantially solved.

Same ambition. A hundred times narrower question.

The difference Not talent, not funding, not data. **One of them had a scoreboard.**

Poll 1 · What actually sank Watson for Oncology?

Commit to one — hands up

Pick the **root** cause, not a symptom.

A

The technology was not good enough yet — 2013 models were too weak.

B

Not enough real patient data was available to train on.

C

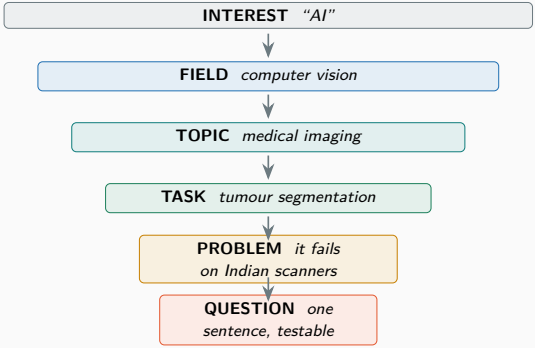
The problem was never narrowed to something with a defined comparison and metric.

D

Doctors resisted it — an adoption and change-management failure.

The narrowing funnel

Six levels from an interest to a question



Where students actually stop.

Level 1–2: "I want to do AI."

Level 3: "AI in healthcare."

Level 5–6: where a project can actually start.

The rule of thumb.

At each level ask: *how many different final-year projects would satisfy this sentence?*

> 100	too broad
5–20	getting there
1	you have a question
0	too narrow — nobody cares

Five knobs — turn any of them to narrow

1 · POPULATION
Who or what is it about?
all text → Hindi–English code-mixed WhatsApp forwards

2 · TASK
Which precise decision?
understand images → segment tumour boundaries

3 · COMPARISON
Against what baseline?
is it good? → better than U-Net and nnU-Net?

4 · METRIC
Measured how, exactly?
accuracy → Dice on the enhancing-tumour class

5 · SETTING
Under which constraint?
unconstrained → one 8 GB GPU, 200 ms per scan

How many knobs should you turn?

Three is usually right. Turn one and you are still too broad. Turn all five and you have engineered away everything anyone would find interesting.

Four more of yours — which knobs did you turn?

Submitted statement	Knobs turned	Knobs still untouched
Mishthi Goyal — "Does a digital fire-safety documentation system improve record management over paper-based records, under district fire office operations?"	Population, comparison, setting	**Metric.** What is better record management — retrieval time? missing-entry rate?
Ishan Kumar — "Does AI-based crop disease detection improve diagnosis accuracy over traditional visual inspection under field conditions?"	Task, comparison, metric, setting	**Population.** Which crop, whose fields, and where do the photographs come from?
Shreyansh Kothari — "Does interval training improve cardiovascular endurance over steady-state cardio, under a 12-week training programme?"	Task, comparison, setting	**Population and metric.** Trained or untrained adults? VO_2max or time-to-exhaustion?
Dhruv Pareek — "Does practising data structures and algorithms improve problem-solving skills of first-year students within 30 days?"	Population, setting	**Comparison and metric.** Better than *what* — and measured on which problem set?

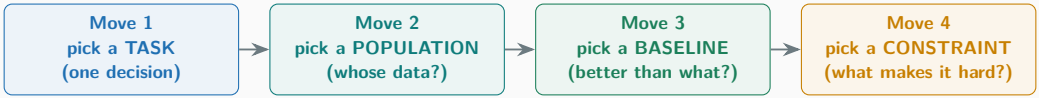
Each of these turned two or three knobs. **One more turn each and they are finished sentences.**

Your turn

Think–Pair–Share · narrow this in four moves

The starting interest

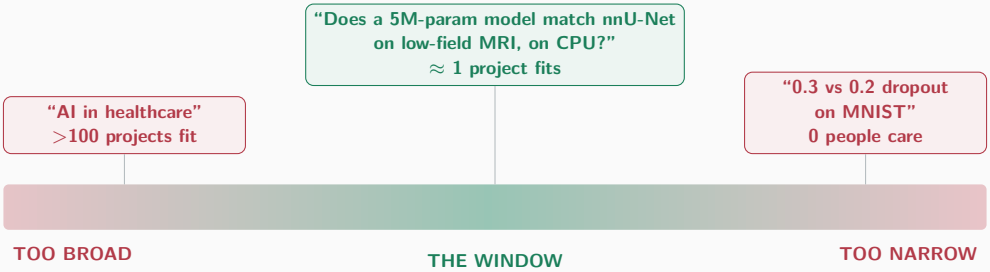
“Artificial intelligence for healthcare.”



Think 90 s · Pair 2 min · Share 3 min

Write the final one-sentence question. Then answer: **how many different projects would satisfy it?**

The window: too broad, just right, too narrow



Symptoms of too broad

You cannot name the dataset. You cannot name the baseline. Your literature search returns 40,000 papers.

Symptoms of too narrow

You can answer it in an afternoon. Nobody outside your lab would read the result. You cannot say who benefits.

Problem-statement templates

Five templates — almost every CS problem fits one

T1 · COMPARATIVE most common

"Does *X* improve *metric* over *baseline*, on *data*, under *constraint*?"

T2 · TRANSFER / POPULATION

"To what extent does *method*, validated on *P1*, retain *metric* on *P2*?"

T3 · EXPLANATORY

"*Why* does *method* fail on *cases*, and which *factor* accounts for the failure?"

T4 · FEASIBILITY / DESIGN

"Can a system achieve *outcome* subject to *C1* and *C2* — and at what cost?"

T5 · MEASUREMENT / VALIDITY

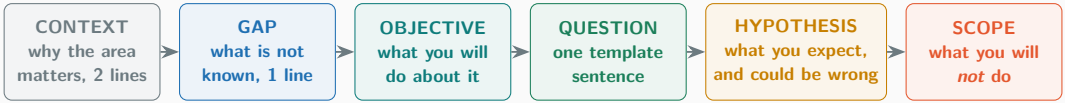
"Does *metric M* actually correlate with *outcome O* that we claim to care about?"

Choosing between them

New method → T1. New setting → T2.
Known failure → T3. Hard constraint → T4.
Suspect the yardstick → T5.

T1 is the default. T3 and T5 are the most under-used — and the easiest to defend, because they need no new system.

The full problem statement — six moves, half a page



The box everyone skips

SCOPE. Writing down what you will *not* do is what stops a supervisor, an examiner — or you at 2 a.m. — from quietly expanding the project.

Length discipline

Six moves. **Under 200 words.** If you need more, you have not narrowed — you have only described.

Worked example · interest to full statement



The written statement — 6 moves, 140 words

[Context] Leaf-disease classifiers reach over 99% accuracy on the PlantVillage benchmark, and several are deployed as advisory apps for smallholder farmers. **[Gap]** PlantVillage images are captured on a uniform background under controlled light; accuracy on photographs taken by farmers in the field is reported far less often, and mechanisms for the drop are not established. **[Objective]** This work measures and explains that degradation for three crops common in Uttarakhand. **[Question]** Does a MobileNetV3 classifier trained on PlantVillage retain top-1 accuracy within 10 points on field-captured images, running offline in under 300 ms on a sub-$80 Android device? **[Hypothesis]** It will not; background clutter will account for more of the loss than illumination. **[Scope]** Three crops only; no multi-disease co-infection; no yield prediction.

Template used: **T2 (transfer)** with a **T3 (explanatory)** second half. Two templates, one project — that is normal.

Checkpoint

Quiz · write all four before any reveal

Q1

Your literature search on your question returns 38,000 papers. This means:

(A) a popular area, good (B) the question is too broad
(C) use better keywords (D) the question is too narrow

Q2

Which is the *scope* statement?

(A) "We will use a transformer" (B) "We evaluate only on Hindi and English"
(C) "Accuracy will improve" (D) "The area is important"

Q3

You suspect the standard metric in your area does not measure what people claim. The right template is:

(A) T1 comparative (B) T2 transfer
(C) T4 feasibility (D) T5 measurement

Q4

Roughly how many distinct final-year projects should satisfy a well-narrowed question?

(A) zero (B) about one
(C) five to twenty (D) as many as possible

Quiz · answers

Q	Ans	Why
1	**B**	The size of your result set is a free diagnostic. C is a fair instinct, but better keywords only help *after* the question itself is narrowed.
2	**B**	Scope says what you will *not* do. A is method, C is hypothesis, D is context.
3	**D**	T5, measurement validity. It needs no new system — which is exactly why it is under-used and easy to defend.
4	**B**	About one. Zero means over-narrowed and irrelevant; five to twenty means you have a topic, not a question.

Score yourself 4/4 excellent · 3/4 solid · $\leq$2 re-read the funnel and the five knobs tonight.

Does your question fit the machine you actually have?

The narrowing you forget. Every constraint you wrote implies a number of experimental runs. Multiply before you commit.

Rule of thumb for a final-year project:

$<$ 40 GPU-hours	comfortable
40–150	plan carefully
$>$ 150	narrow another knob

The cheapest knob to turn is usually the number of baselines or the number of seeds — not the science.

I.4 turns this into a formal feasibility criterion. Today it is just arithmetic.

```python
# Cost of the experiment your question IMPLIES.
# Run this BEFORE you commit to the statement.

methods = 2 # yours + how many baselines?
datasets = 2 # PlantVillage + field photos
seeds = 3 # minimum for any honest claim
ablations = 4 # background, lighting, blur, resolution
hours_per_run = 1.5

runs = methods * datasets * seeds * (1 + ablations)
total = runs * hours_per_run
print(runs, "runs |", total, "GPU-hours")
# -> 60 runs | 90.0 GPU-hours ... feasible, but tight

# Now add ONE more baseline you were tempted by:
methods = 3
print(methods*datasets*seeds*(1+ablations)*hours_per_run)
# -> 135.0 GPU-hours ... and one more dataset -> 202.5
```

Scope creep is multiplicative, not additive. One extra baseline cost you 45 hours.

Guided practice · rewrite the statement you brought



Individual · 6 minutes · then 2 volunteers read theirs aloud

Use the statement you submitted after I.2. If you have three, pick the one you would actually want to spend a year on.

Stuck at step 2? Turn the **population** knob first — it is the one nobody else will have turned the same way you do.

Close

Reflection and key takeaways

Answer these in your head

1. What did we learn?

2. Was your statement too broad, or too narrow?

3. Which knob will you turn tonight — and what will it cost you in runs?

1 · Narrowing is a funnel, not a hunch

Interest → field → topic → task → problem → question. Six levels, repeatable.

2 · There is a window, with two walls

Too broad wastes months. Too narrow produces a finished project nobody reads. Aim for "about one project fits".

3 · Five knobs, turn about three

Population · task · comparison · metric · setting.

4 · Write it in a template, and cost it

T1–T5, six moves, under 200 words — then multiply the runs before you commit.

Homework, reading and what comes next

Homework — due before Lecture I.4

H1. Your problem statement in **six moves, under 200 words**. Name the template (T1–T5).

H2. The count: how many distinct final-year projects satisfy your question? Justify in one line.

H3. Cost it: methods × datasets × seeds × ablations × hours. Does it fit?

H4. Your **scope** line: three things you will *not* do.

Further reading

Textbook. Melville & Goddard, Ch. 2–3. Ganesan — scope. **Papers.**

Jumper et al., *AlphaFold*, Nature 2021 — first page only; notice how narrow the stated task is.

Strickland, *How IBM Watson Overpromised and Underdelivered on AI Health Care*, IEEE Spectrum 2019.

Wobbrock & Kientz, *Research Contributions in HCI*, 2016 — seven contribution types.

Bring to I.4

Your 200-word statement. We test it against formal criteria — and some will not survive.

Next · Lecture I.4 *Criteria and characteristics of a good research problem* — novelty, feasibility, relevance, measurability, ethics.

“If I had an hour to solve a problem, I would spend fifty-five minutes defining it.”

— widely attributed to Albert Einstein