BenefitLens

An End-to-End Data Pipeline for
Personalized Student Support Discovery and Application Drafting

장신형, 김혜빈, 서재은

Data Engineering, Fall 2026

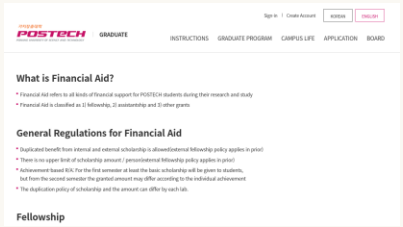
Support beyond familiar notices

85.4% concerned about rising living costs

45.3% say rising prices may affect whether they keep studying

Where can students find support?

University



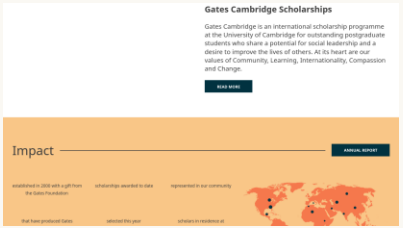
POSTECH financial aid

Company



Google PhD Fellowship

Outside Korea



Gates Cambridge

Government



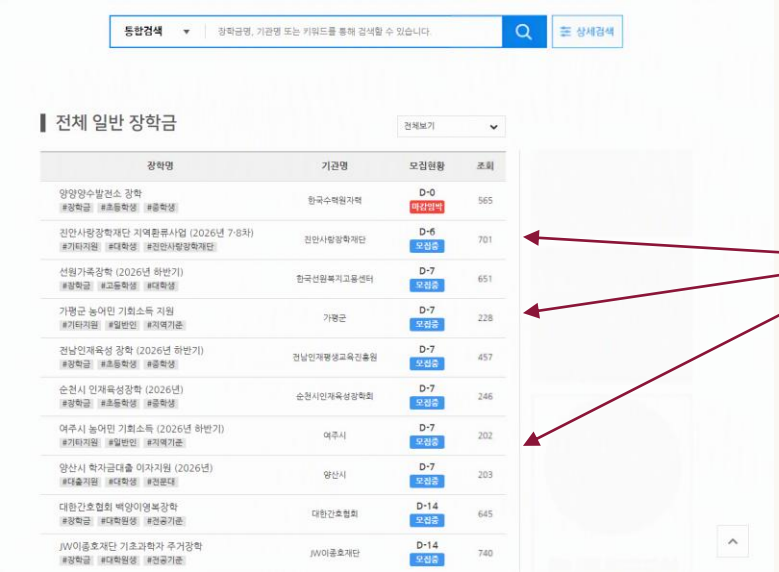
Graduate Presidential Science

Which opportunities fit this student?

Nature 2022 survey, public responses N = 3,253. doi:10.6084/m9.figshare.21277575

Checking fit takes repeated work

Every notice needs a new eligibility check



Student information	Example
Degree / status	PhD / enrolled
Coursework	Incomplete
Support sought	Tuition / stipend

Degree? Dates? Other conditions?

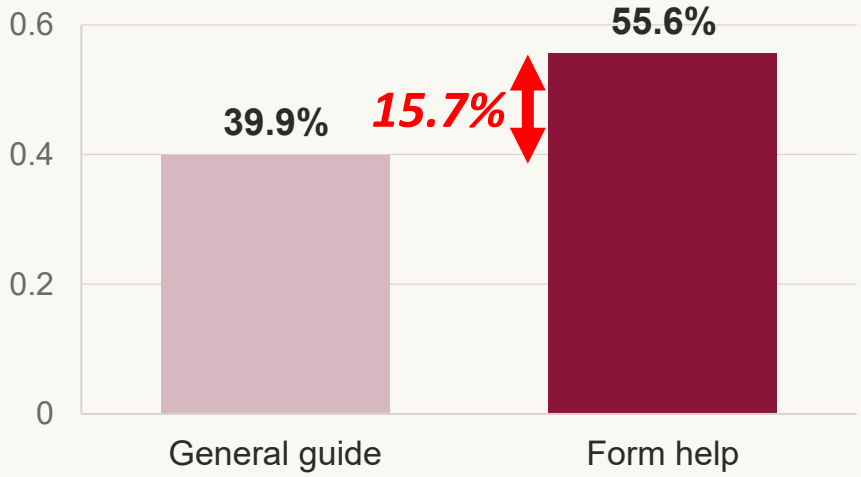
Dreamspon. (n.d.). General scholarship list. Retrieved October 19, 2026, from Dreamspon General Scholarships

Finding support still leaves paperwork

FAFSA (Free Application for Federal Student Aid) study :
868 US students in high school seniors from lower-income families

Group	What participants received
General guide	College costs and aid guidance
Form help	Aid information plus help filling in and sending the form

Students who sent in the form



Students also need help preparing the application.

Bettinger et al., QJE 2012, Table III. Rates use rounded control mean + treatment effects.

What should BenefitLens provide?

Matched notices and help with the first draft

01 Matched notices

Notice	What the student gets
Program and round	Official source and form
Eligibility	Matches and missing facts
Funding and dates	Amount and deadline

02 Initial writing outline

Draft structure	Applicant facts
Problem / goal	Research question
Contribution	Own work and evidence
Next steps	Research plan

Goal: Easier discovery and application preparation

Prepared data for student requests

System Preparation

Collect sources

Notices, forms and writing examples

Extract information

Common fields and writing structures

Store in the DB

Linked records for later requests

When a student asks

Student input

Profile and application facts

Use stored records

Match conditions and prepare an outline

Return the results

Official notice, outline and writing guidance

Search expressions expand the crawl

System Preparation

Each round collects sources and expands the search dictionary with newly discovered expressions

The process repeats until no new expressions are found in two consecutive rounds

Starting dictionary

studentship
tuition waiver
research statement

New source phrases

assistantship
Impact Statement

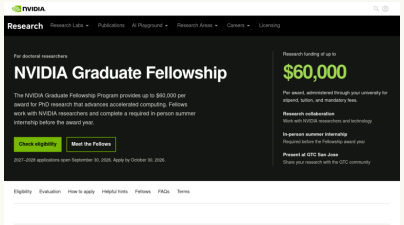
Next searches

Add verified terms.
Search again.

Original material to structured fields

System Preparation

01
Original material



Official HTML

02
Text extraction

active PhD student

up to $60,000
per award

Apply by
October 30, 2026.

Text parser / OCR

03
Field extraction

LLM

Degree: PhD

Max. USD 60,000

Due: 30 Oct 2026

04
DB schema

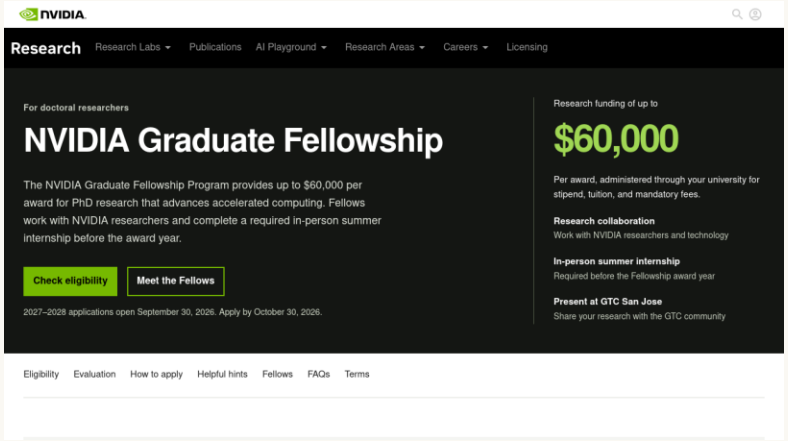
Field	Type
eligibility	JSONB
benefits	JSONB
application	JSONB
source_refs	JSONB

Each stored value retains its source evidence.

An official notice in the common schema

System Preparation

NVIDIA Graduate Fellowship



One program, one recruitment round

Column / JSON path	Stored value
round_id	2027-2028
application.closes_on	2026-10-30
benefits[0]	Maximum USD 60,000
eligibility.all	PhD and enrolled
additional_checks	First PhD year, field and other checks
application.prompts	Proposal: 2 pages, CV, 2–3 letters

Source: NVIDIA Graduate Fellowship, https://research.nvidia.com/graduate-fellowships

Finding public application materials

System Preparation

How can applicants prepare these documents?

Search and expand

Program + applicant experience terms

NVIDIA Fellowship
proposal / research statement
application experience / interview

Collect the original sources


Performance Accessibility Through Distributed Array Programming

Rohan Yadav
Stanford University

Rohan Yadav
Public proposal



Fellowship panel
Preparation accounts

NVIDIA: 3 applicant cases, 5 documents

Writing structures linked to the notice

System Preparation

The LLM reads full source texts and extracts writing structures.



Research Summary

Yuke Wang

5 full source texts
Case IDs retained

LLM

program_knowledge	Stored writing knowledge
writing_pattern[]	Problem / goal, contribution, next steps
required_inputs (in each pattern)	Applicant's problem, own evidence and plan
experience_cases[]	Case IDs and source quotes

Student profile and writing facts

When a student asks

1 Basic profile

Degree: PhD

Status: enrolled

Coursework: incomplete

Support: tuition / stipend

2 Compare conditions

NVIDIA condition	Action
PhD / enrolled	Match basic profile
First PhD year	Confirm year / review exception
In-person internship	Confirm availability

3 Applicant's writing facts

Official prompt
+ writing structure

Applicant's goal, own contribution and evidence

Initial outline
+ follow-up questions

Conclusion

Problem	BenefitLens response
Notices are scattered	Collect originals and store common fields
Eligibility checks repeat	Match profiles and ask for missing facts
Finding a notice still leaves writing	Link official prompts with writing structures

Students receive matched notices and a starting outline.

BenefitLens

Thank You

Questions?