Survey introduction: This survey is about the possible use of predictive models to help the university identify and contact students who may be experiencing academic difficulties, so that support can be offered early. The goal of these questions is not to evaluate any individual student, and there are no right or wrong answers. We are interested in your perspective on how such a system should be designed and used, and under what conditions it would be acceptable.

In this study, predictive models are considered as tools to help the university identify **students who may be experiencing academic difficulties**, so that support can be offered proactively.

The model would be trained using patterns from past student outcomes, such as DFW rates (students earning a D or F, or withdrawing from a course), because these are among the few outcomes the university can reliably observe. However, the purpose of the system is not only to predict DFW outcomes, but more broadly to support earlier and more targeted outreach to students who may be experiencing academic difficulties.

Throughout the survey, please assume that any data used follows applicable privacy laws and institutional safeguards, and that the model is used only to offer support to students.

Q4.1 **Imagine the university uses a predictive model to identify students who may be experiencing academic difficulties so it can offer support early.**

The model begins with course credits earned. The university could add other data sources that may improve identification but may raise privacy concerns.

There are competing perspectives:
Some argue we should use any data that helps identify students experiencing academic difficulties - not using available data means missing students who need help.
Others argue that some data is too invasive - surveillance risks and privacy violations can outweigh the benefit of identifying more students experiencing academic difficulties.

In each of the questions below, please consider whether it would be acceptable for the model to **use the additional type of data described**, given the data it already uses at that point.

In the next questions, an ‘improvement’ means: the model correctly identifies more students who are actually experiencing academic difficulties **out of 100**, while the number of students incorrectly flagged stays about the same (incorreclty flagged means incorrectly identfied as experiencing academic difficulties).

Q4.2 Would it be acceptable for the model to additionally use **course performance data** (such as grades, passed or failed courses, or number of registered courses) in addition to a model that uses only the number of course credits earned?

- ○ Acceptable even if it does not improve identification
- ○ Acceptable only if there is at least a small improvement (about 1–3 additional students identified out of 100 students experiencing academic difficulties)
- ○ Acceptable only if there is at least a moderate improvement (about 5 additional students identified out of 100 students experiencing academic difficulties)
- ○ Acceptable only if there is at least a large improvement (about 10 additional students identified out of 100 students experiencing academic difficulties)
- ○ Never acceptable

Q4.3 Would it be acceptable for the model to incorporate **administrative admissions and enrolment data** (e.g., high-school GPA, program, school/college), in addition to course credits earned and course performance metrics?

- ○ Acceptable even if it does not improve identification
- ○ Acceptable only if there is at least a small improvement (about 1–3 additional students identified out of 100 students experiencing academic difficulties)
- ○ Acceptable only if there is at least a moderate improvement (about 5 additional students identified out of 100 students experiencing academic difficulties)
- ○ Acceptable only if there is at least a large improvement (about 10 additional students identified out of 100 students experiencing academic difficulties)
- ○ Never acceptable

Q4.4 Would it be acceptable for the model to also use **online course activity data from the university learning platform** (for example, how often course pages are accessed or whether assignments are submitted late) in addition to a model that uses course credits earned, course performance data, and admission and enrolment data?

- ○ Acceptable even if it does not improve identification
- ○ Acceptable only if there is at least a small improvement (about 1–3 additional students identified out of 100 students experiencing academic difficulties)
- ○ Acceptable only if there is at least a moderate improvement (about 5 additional students identified out of 100 students experiencing academic difficulties)
- ○ Acceptable only if there is at least a large improvement (about 10 additional students identified out of 100 students experiencing academic difficulties)
- ○ Never acceptable

Q4.5 Some types of personal data are often considered particularly sensitive and raise additional privacy or ethical concerns.

In this question, each sensitive data type is considered as an **additional data source**, added on top of all non-sensitive data described earlier in this section.

For each type of sensitive data below, indicate the **minimum improvement** that would make it acceptable to include this data in the model as in the questions above.

Again, an improvement means correctly identifying additional students who are actually experiencing academic difficulties, out of 100 such students, while keeping the number of students incorrectly flagged at about the same level. A small improvement: about 1–3 additional students A moderate improvement: about 5 additional students A large improvement: about 10 additional students

	Acceptable if it does not worsen the model	Acceptable only if there is at least a small improvement	Acceptable only if there is at least a moderate improvement	Acceptable only if there is at least a large improvement	Never acceptable
Age	○	○	○	○	○
Gender	○	○	○	○	○
Nationality	○	○	○	○	○
Disability-related information	○	○	○	○	○
Race/ethnicity	○	○	○	○	○
First-generation college student status	○	○	○	○	○
Financial aid status	○	○	○	○	○

Q4.6 Sometimes, including sensitive personal data does not increase how many students experiencing academic difficulties are identified overall, but it may help improve the fairness of the model.

In this context, improving fairness means reducing differences across student groups in how well the model identifies students experiencing academic difficulties or avoids unnecessarily flagging students who are not experiencing difficulties.

For each type of sensitive data below, please indicate the minimum condition under which you would find its use acceptable in **situations where it does not increase the total number of students identified**.

	Acceptable even if it does not improve overall identification or fairness	Acceptable only if it improves fairness across student groups	Never acceptable	Not sure
Age	○	○	○	○
Gender	○	○	○	○
Nationality	○	○	○	○
Disability-related information	○	○	○	○
Race/ethnicity	○	○	○	○
First-generation college student status	○	○	○	○
Financial aid status	○	○	○	○

Q5.1 Predictive models for identifying students **who may be experiencing academic difficulties** differ in how clearly their decisions can be explained.

In this question, explainability refers to how well it is possible to explain **why a specific student was identified or not identified** as potentially experiencing academic difficulties.

There are competing perspectives:

- Some argue that explainability is essential. Counselors and students need to understand why someone was flagged to build trust and provide appropriate support, even if it means identifying fewer students who are experiencing academic difficulties.
- Others argue that effectiveness matters most. If a less explainable model identifies more students who are experiencing academic difficulties, that benefit outweighs the difficulty in explaining decisions.

Both perspectives raise important questions. The following questions ask about your views on explainability trade-offs.

For each level of explainability below, indicate the minimum improvement that would make it acceptable to use a model with that level of explainability, compared to a higher level of explainability on this scale.

In the questions below, an improvement means correctly identifying additional students who are actually experiencing academic difficulties, out of 100 such students, while keeping the number of students incorrectly flagged at about the same level.

A small improvement: about 1–3 additional students
A moderate improvement: about 5 additional students
A large improvement: about 10 additional students

Levels of explainability (with examples)

Very high explainability (single clear rule)
Each decision is based on one simple and transparent rule. Example: a student is identified if they have passed fewer than a certain number of course credits.

High explainability (clear explanation)
Each decision can be explained using a small set of clear reasons with stable direction. Example: a student is identified because of a combination of low course credits, recent failed courses, and delayed progression, and it is possible to explain how each factor contributed.

Moderate explainability (partial but informative explanation) Each decision can be explained in general terms by highlighting the most influential factors, but the full reasoning cannot be reduced to a small, stable explanation. Example: A student is identified based on many aspects of their study history, where it is possible to point to the most important contributing factors, but not fully explain the decision in a simple way.

Low explainability (limited, case-specific explanations) Only limited or approximate explanations are possible. For a specific student, it may be possible to give a rough or partial explanation (for example, highlighting some contributing factors), but the explanation is incomplete and difficult to communicate clearly.

Very low explainability (no meaningful case-specific explanation) Decisions are produced by highly complex models whose internal reasoning cannot be meaningfully interpreted. For a specific student, it is generally not possible to explain why they were identified or not identified, beyond stating that the model detected a pattern.

	Acceptable even if it does not improve identification	Acceptable only if there is at least a small improvement	Acceptable only if there is at least a moderate improvement	Acceptable only if there is at least a large improvement	Never acceptable
High explainability (few clearly defined factors)	○	○	○	○	○
Moderate explainability (many additive factors)	○	○	○	○	○
Low explainability (limited, case-specific explanations)	○	○	○	○	○
Very low explainability (no meaningful case-specific explanation)	○	○	○	○	○

Q6.1 When a predictive model is used to identify students who may be experiencing academic difficulties, it may not work equally well for all student groups. For example, the model might correctly identify 80% of students experiencing academic difficulties in one group but only 65% in another group. Reducing these differences across student groups may require accepting that fewer students are correctly identified overall. For instance, making the model work equally well for all groups might mean it now identifies 70% of students experiencing academic difficulties in all groups, helping some groups more but helping other groups less than before.

There are competing perspectives:

- **Some argue** that overall effectiveness matters most. If more total students are helped, the model is doing its job even if it works better for some groups than others.
- **Others argue** that unequal performance perpetuates discrimination. A model that works worse for already-marginalized students deepens existing inequalities rather than helping.

Both perspectives raise important questions. The following questions ask about your views on fairness trade-offs.

In this question, consider situations where changes to the model make its performance more equal across student groups, but reduce the total number of students who are correctly identified early enough to offer support.

How large of a decrease in correctly identifying students who are experiencing academic difficulties would you find acceptable in order to make the model work equally well across different student groups?

- ○ No decrease is acceptable, even if this means the model does not work equally well across student groups
- ○ About 1–3 fewer students identified out of 100 students experiencing academic difficulties
- ○ About 5 fewer students identified out of 100 students experiencing academic difficulties
- ○ About 10 fewer students identified out of 100 students experiencing academic difficulties
- ○ Not sure

Q6.2 When a predictive model is used to identify students who may be experiencing academic difficulties, it may make different types of errors for different student groups. For example, the model may be more likely to miss students who need support in some groups, or more likely to unnecessarily contact students who do not need support in other groups. When deciding how

to make the model fair across student groups, which of the following should be prioritized to be equal across groups?

- ○ Making the model equally good across student groups at identifying students experiencing academic difficulties and would otherwise be overlooked by the system
- ○ Making the model equally good across student groups at avoiding unnecessary contact with students who are not experiencing academic difficulties
- ○ Both goals should be prioritized equally
- ○ Not sure

Q7.1 If a model indicates that a student may be experiencing academic difficulties, the university could respond in different ways. Which of the following interventions do you think would be most acceptable? **(Please select up to two.)**

- [ ] Receiving a confidential email or message with general information about available support services
- [ ] Receiving a confidential, personalized email or message that refers to the student's situation and suggests relevant support options
- [ ] An invitation to a voluntary meeting or call with a student advisor
- [ ] Receiving a confidential phone call from a student advisor to discuss support options
- [ ] No proactive outreach is taken; students receive support only if they explicitly request help
- [ ] Other (please fill in) _____

Q7.2 How often should such outreach occur?

- ○ Once per year
- ○ Twice per year
- ○ Whenever the model is updated (for example, after new course results or progress updates)
- ○ Not sure

Q8.1 When a system is used to identify students **who may be experiencing academic difficulties**, it has to decide how selective it should be.

Assume that students identified by the system receive **a brief email with general information about available support services.**

One approach is to contact many students, even if some of them turn out not to need support. Another approach is to contact fewer students, which reduces unnecessary contact but may miss some students experiencing academic difficulties. Which approach do you think the university should take in this case?

- ○ Contact many students, even if some do not actually need support
- ○ Slightly prefer contacting many students, accepting that some students who do not need support may be contacted
- ○ Aim for a balance between contacting many students and contacting fewer students
- ○ Slightly prefer contacting fewer students, accepting that some students experiencing academic difficulties may be missed
- ○ Contact fewer students, even if some students experiencing academic difficulties are missed
- ○ Not sure

Q8.2 Now assume that students identified by the system receive **a phone call or personalized email or message that refers to the student's situation and suggests relevant support options** from a student advisor. In this case, which approach do you think the university should take?

- ○ Contact many students, even if some do not actually need support
- ○ Slightly prefer contacting many students, accepting that some students who do not need support may be contacted
- ○ Aim for a balance between contacting many students and contacting fewer students
- ○ Slightly prefer contacting fewer students, accepting that some students experiencing academic difficulties may be missed
- ○ Contact fewer students, even if some students experiencing academic difficulties are missed
- ○ Not sure

Q9.1 Overall, which of the following best reflects your view on the development and use of predictive models for identifying students who may be experiencing academic difficulties?

- ○ I would generally support their use
- ○ I would support their use only if relevant experts (such as academic, technical, and ethics or governance experts) are meaningfully involved in the design and oversight of the system
- ○ I would support their use only if students and advisors are meaningfully involved alongside experts in the design and oversight of the system
- ○ I would support their use only if students and advisors are meaningfully involved in the design and oversight of the system
- ○ I would not support their use

Q9.2 Is there any important concern, consideration, or perspective related to the use of predictive models for student support that you feel was not covered in this survey?

_____

_____

_____

_____

_____