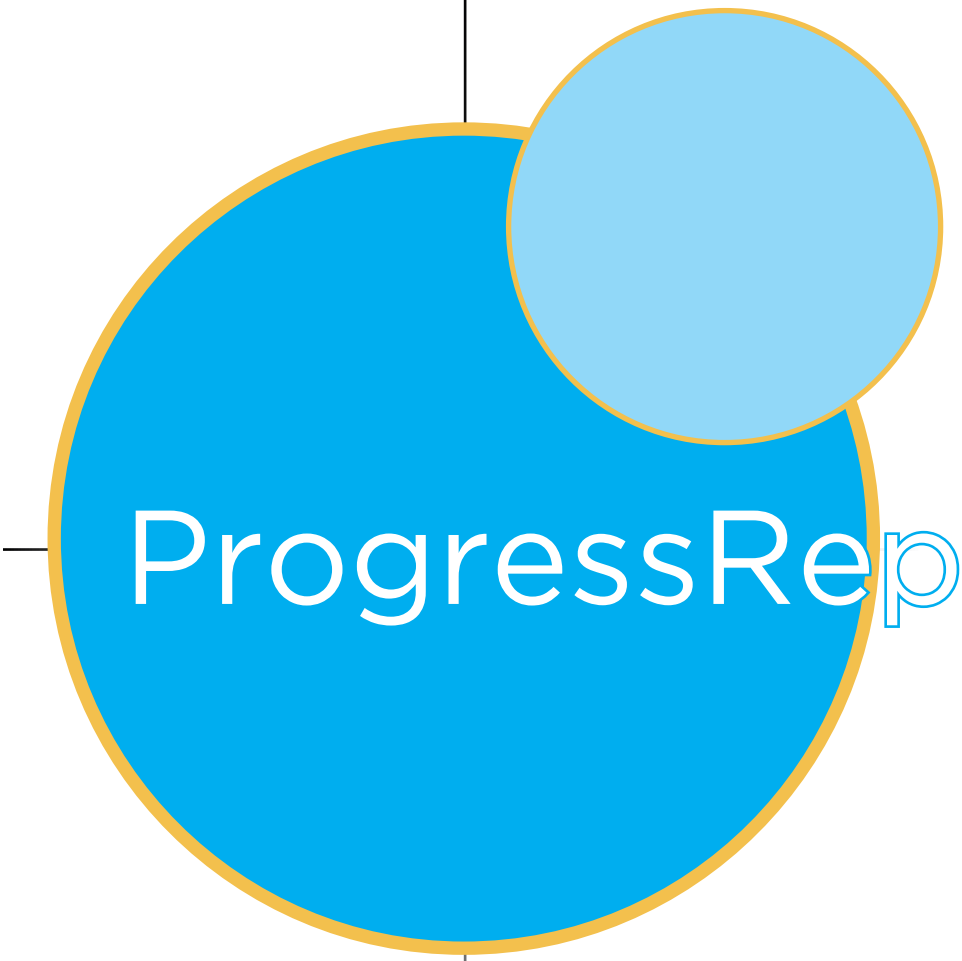

ProgressReport

TABLE OF CONTENTS

Users Guide

2 What We Do...
5 Login + Account Creation
6 Inputting your Courses
9 My Board: Keeping Track of Interested Majors and Certificates
13 Suggested Majors/Certificates
17 Suggesting Courses
18 Tracking Requirements

Developers Guide

23 Introduction
23 Back-End Systems
28 Transcript Parsing
29 Progress Structure in Middle-Tier
29 Course Suggestion Algorithm
30 Front-End Overview
31 CONCLUSION

Users Guide

Firstly, thank you for taking an interest in Progress Report! We hope that you will find our product useful in your endeavors, but first, we'd like to tell you about what we want to help you accomplish with our product.

What We Do...

Deciding what to major in is often a difficult and stress-filled decision for students, as it not only impacts the remaining three to four years of one's time at Princeton, but also where and what one will be doing in the future. Even when students decide their major, there is the complicated mire of degree requirements that students have to wade through in order to ensure graduation.

Once you add certificates on top of one's major, keeping track of what classes are fulfilling which particular major requirement, or which classes can fulfill both a major and a certificate requirement, or what classes one still has to take in order to graduate becomes an extremely complicated process.

What We Do... (Cont.)

Currently, the only mechanism in place for students to determine which distribution and degree requirements that they are fulfilling is TigerHub. However, TigerHub only provides the most important aspects of this information after students have declared a major, which, for many, is after their fourth semester at Princeton.

Knowing what classes you have already taken prior to declaring your major can often be useful in the major decision process, since you can see whether courses you have already taken fulfill the major prerequisites or requirements.

Furthermore, finding a single, centralized source of all major requirements is challenging, as the course registrar page is often outdated, leaving searching through each department's website as the only alternative for students who wish to compare different major requirements.

What we hope to accomplish with ProgressReport is to give students an easy-to-use, informative platform to determine what major and certificate requirements they have fulfilled, whether they are seniors trying to ensure that they have

What We Do... (Cont.)

taken all the necessary courses to graduate, freshmen trying to decide what they want to major in, or sophomores deciding what certificates they might pursue.

There are countless ways for students to make use of our platform, some of which we will cover in this guide, and some that we don't know of yet, but hope that you will discover through your academic journey here at Princeton.

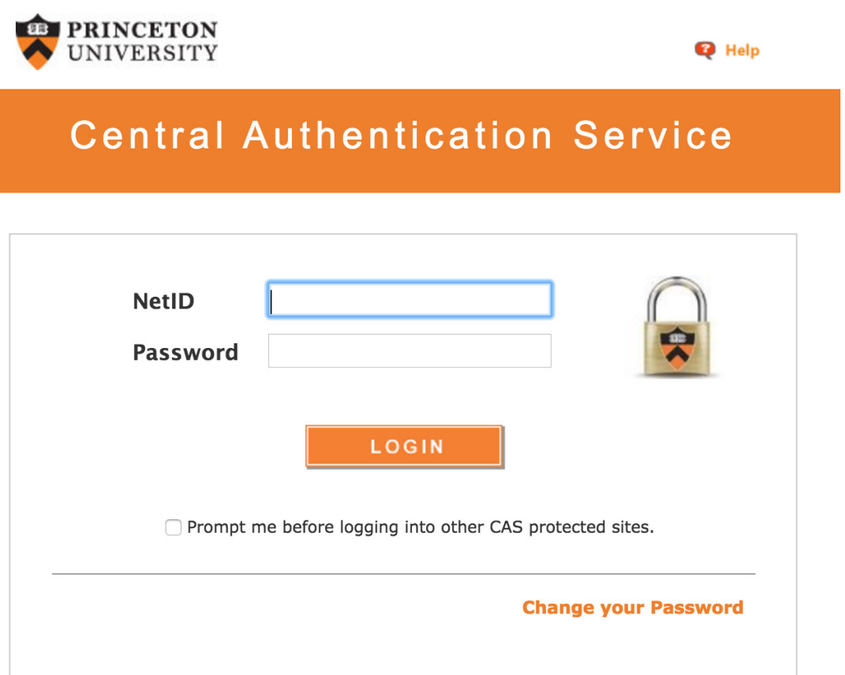
Throughout this guide, aside from the main information, there will be tips in the sidebars regarding different possible ways to use the features taken from advice from beta-tester students.

Now that we've gotten through our introduction, let's take a look at how to use ProgressReport!

FIRST STEPS

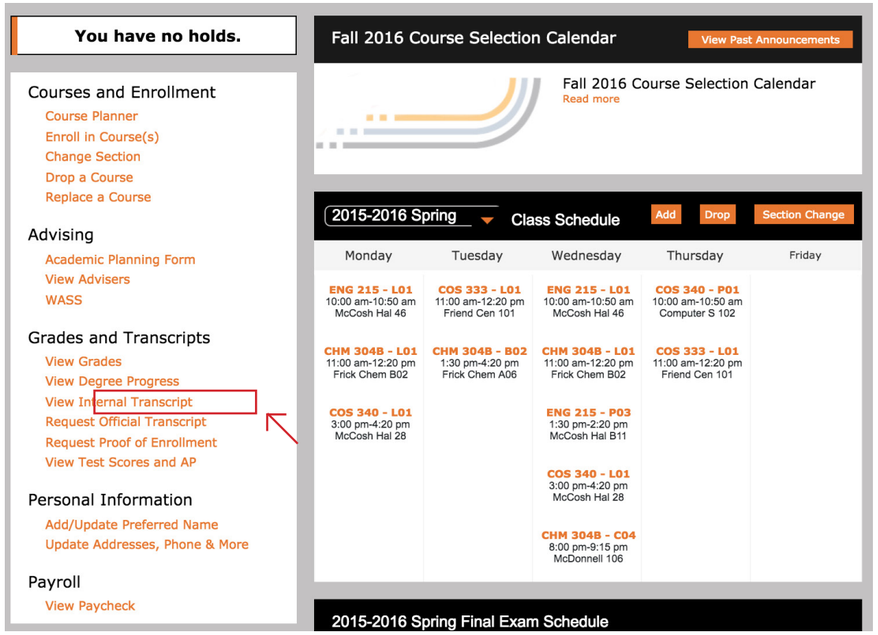
Login + Account Creation

Our login system is integrated with the Princeton Central Authentication System (CAS), which means that you don't need to create any extra account to use ProgressReport. When you first reach the page at http://progressreport.herokuapp.com/ you will be redirected to the login page for CAS, which looks like:



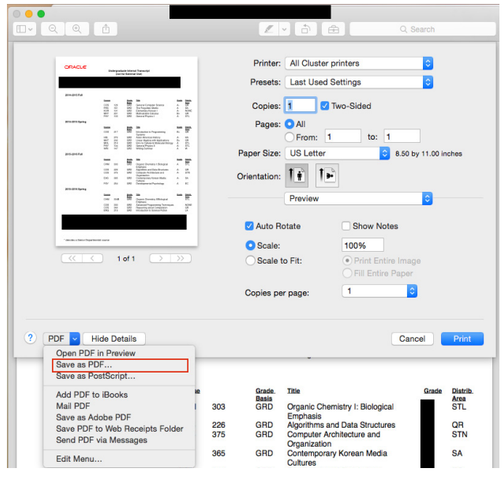
Inputting Courses

After logging in, you will be taken to back to ProgressReport and be asked to upload your transcript or to input courses manually. The transcript that we ask you to upload is the internal TigerHub transcript that students can request immediately by logging into TigerHub and clicking on the View Internal Transcript button.



Inputting Courses (Cont.)

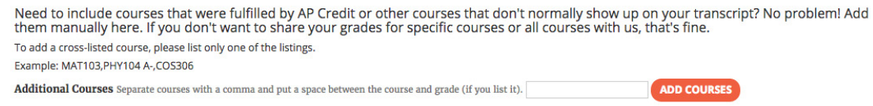
After downloading your transcript, you need to decrypt it in order for our website to be able to retrieve your course information from your transcript. The decryption process is simple. For MAC users, you can open the pdf file in Preview, and select to print it. When the Printing window pops up, select "Save as PDF" which will save a decrypted version of your transcript on your computer. For PC users, we have further instructions on the website itself to aid you in the decryption process.



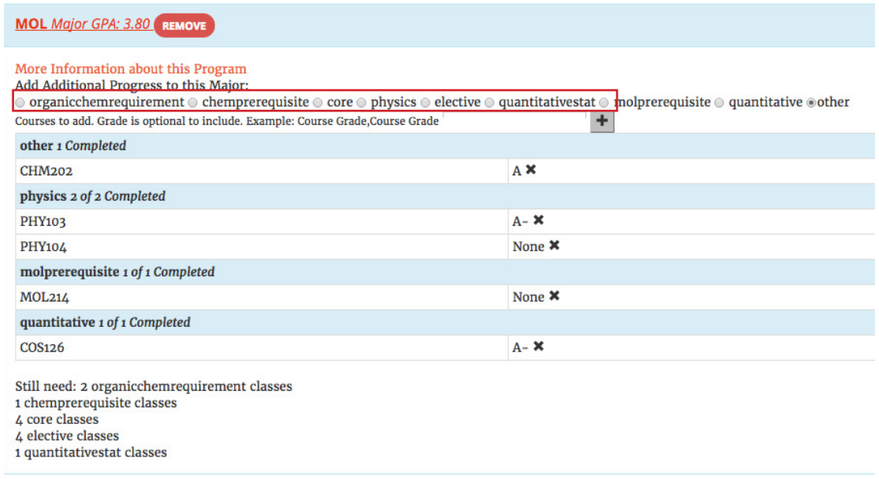
After decrypting your transcript, all you have to do is upload that copy to ProgressReport by clicking on the Upload Transcript button.

Inputting Courses (Cont.)

You can also input courses manually into the user input box which is below the Upload Transcript button.



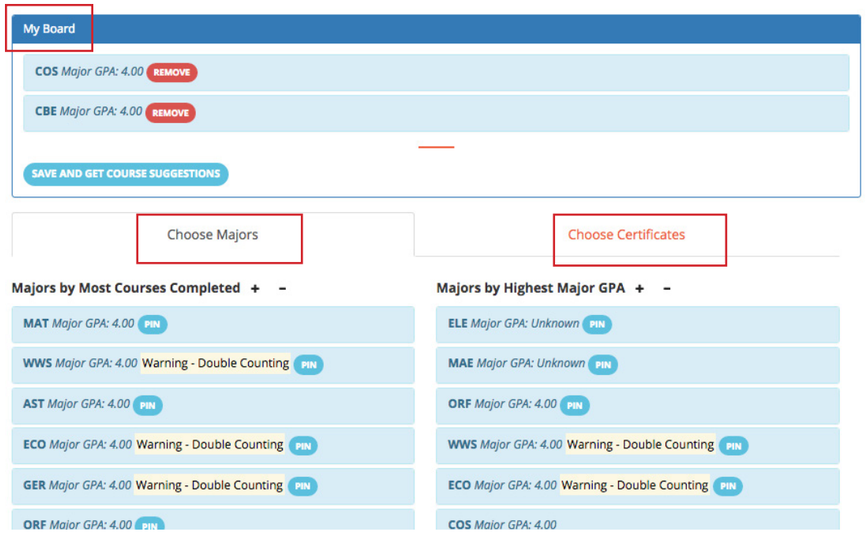
You can also decide to only manually add specific courses to a single major or certificate of interest by selecting a major or certificate and using the Manual Course addition feature under each major and certificate list.



FEATURES

My Board: Keeping track of Interested Majors and Certificates

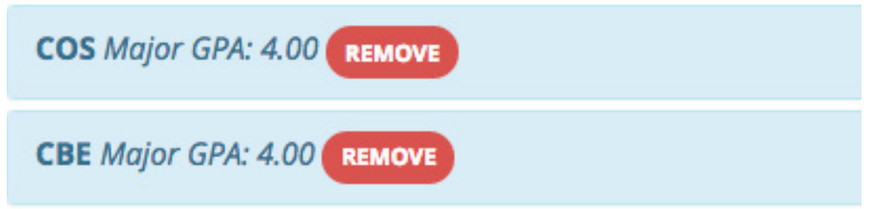
Now that all of your classes have been uploaded, you should note that the page consists of a My Board section, and a Choose Majors and Choose Certificates section.



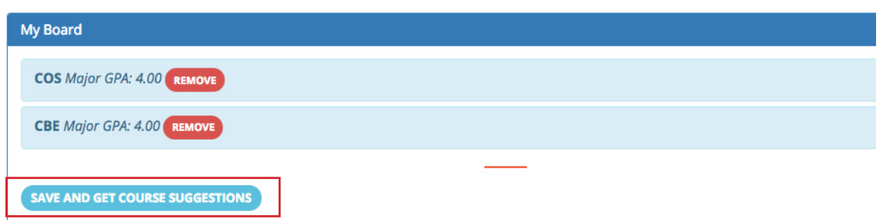
You can pin the Majors and Certificates which you are most interested in for easy access, which will add it to the My Board section for future visits to our site.

My Board (cont.)

Should you decide that you are no longer interested in a major, you can click the remove button next to it to remove it from your board.



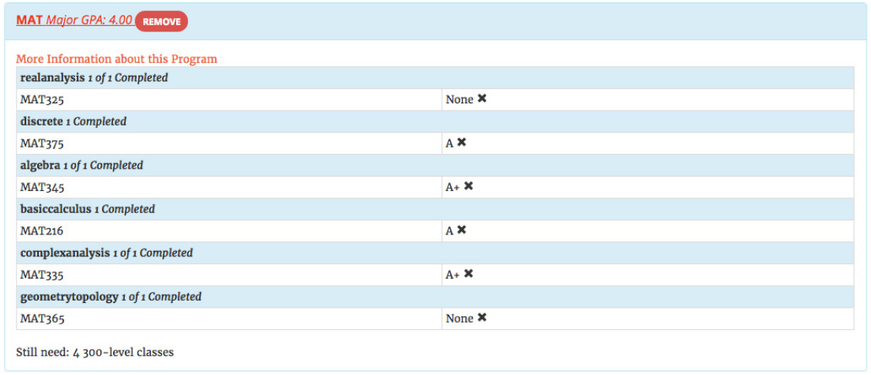
After making any changes to your board, whether it is pinning new majors/certificates or removing them, make sure to click the SAVE button to ensure that your changes will be kept after you logout or refresh the page. This SAVE button will also generate a list of recommended courses based on your chosen majors and certificates of interest.



My Board (Cont.)

If you want to see which requirements you have fulfilled for a particular major/certificate, all you have to do is click on the major/certificate and a list will expand to show you which requirements you have fulfilled, as well as your grades for those classes.

You will also be able to see how what you still need to complete in terms of the number of courses required for specific tracks of the major or certificate.



How can these features be used?

Being able to pin majors and certificates to your Board is meant to allow you easier, continued

My Board (cont.)

access to majors that you are especially interested in. Having all this information in a single location means that you can compare two majors or certificates side by side which can help in your major selection, if you are still undecided between several choices. It can also help by showing if you meet the prerequisites necessary for a major, and what electives you have already fulfilled, which may also be a significant factor in your decision-making processes.

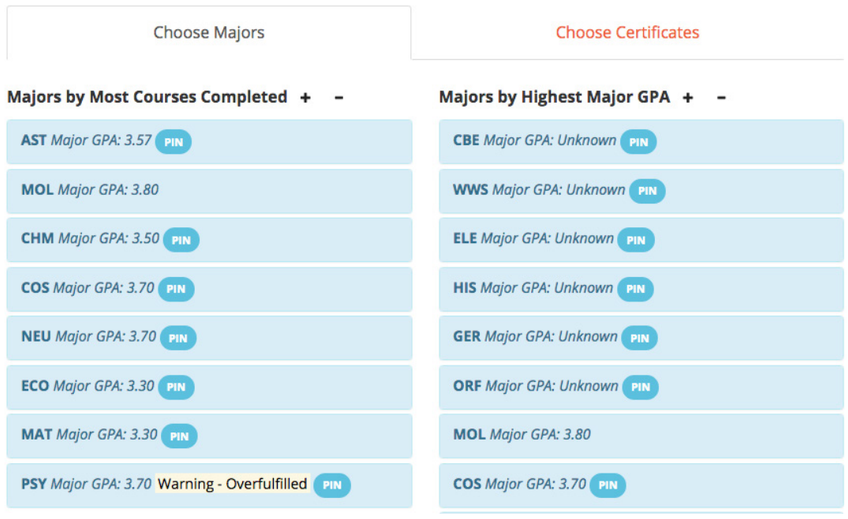
Another way that having the Board is important, even for those who have declared their majors, is the easy accessibility of seeing certificate requirements. For many students, there may be certificates that they can acquire with little extra work. Having such certificates pinned to your Board means that you can see whether classes that you have already taken fulfill some of the courses required for certificates.

Additionally, for those who have committed to their majors and certificates, it allows you to continually monitor your progress in your major. In contrast with TigerHub's degree progress tool, your completed requirements are presented in a more clear manner.

We Suggest Majors/Certificates

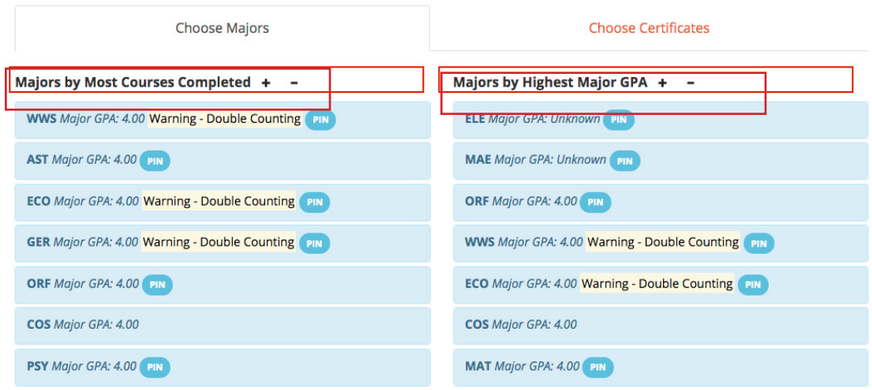
As we mentioned in the previous section, My Board section allows students to keep track of particular majors and certificates they are interested in. But what if you don't really know what you are interested in? And how do you "pin" things to your board? Well, the Choose Majors and Choose Certificates Tabs are there for that purpose!

Below My Board is the Choose Majors and Choose Certificate Tabs. The default selection for these tabs is to be on the Choose Majors tab.



Suggesting Majors/Certificates (cont.)

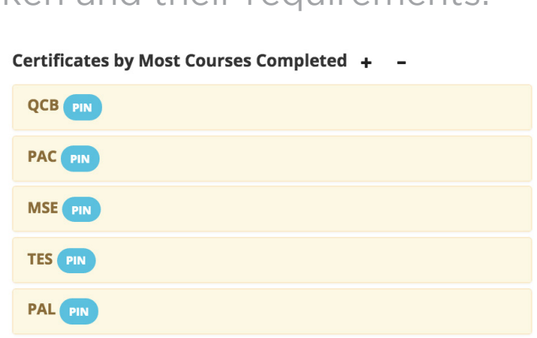
In the Choose Majors tab, there are two lists, one which lists suggested majors for you based on how many requirements you have completed in the major, another that lists suggested majors by your major GPA.



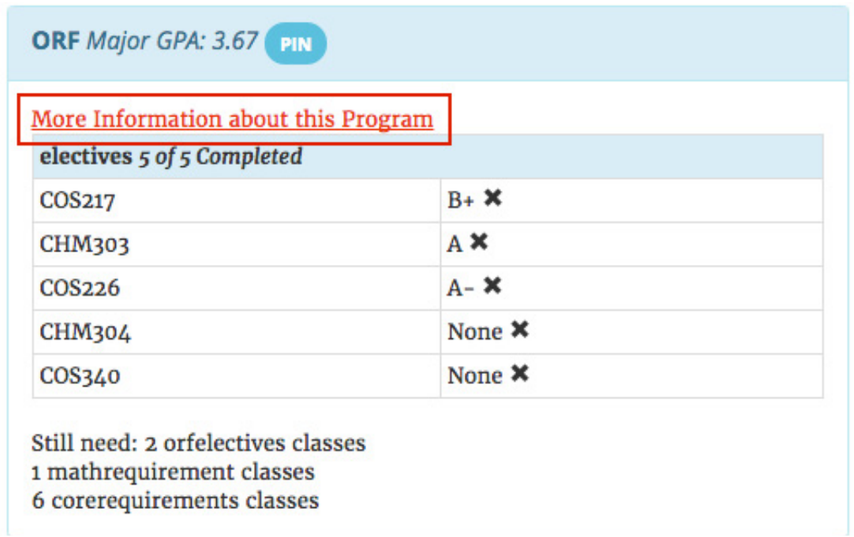
Like with My Board, if you want to see exactly which requirements you have fulfilled for the majors or certificates in the Choose Majors/Certificates tabs, you can just click on the major/certificate to see a list of requirements that you have filled.

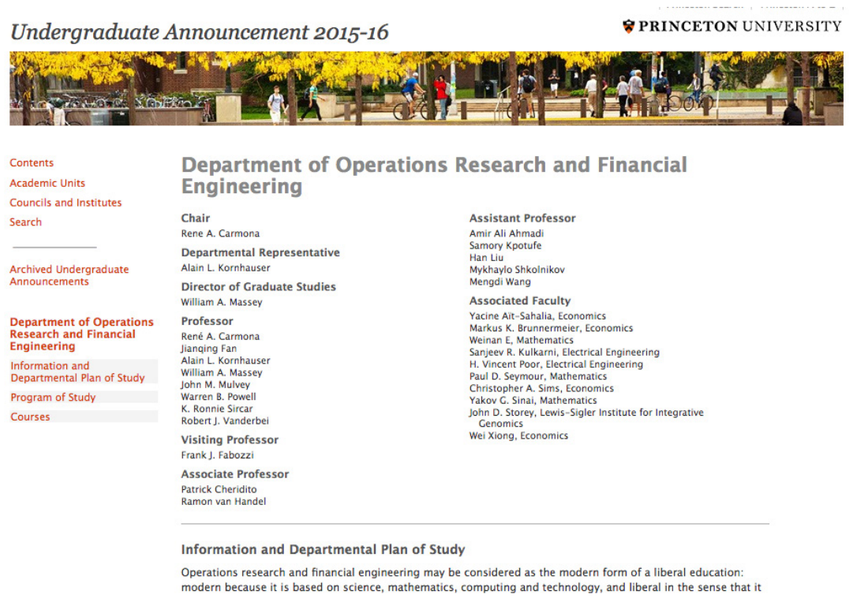
Suggesting Majors/Certificates (Cont.)

Similarly, if you want to look at suggested certificates, click on the Choose Certificate tab. Like with the Choose Majors tab, there is a list of Certificates by Most Courses Completed which gives a list of certificates that have the most overlap between the courses that you have already taken and their requirements.



Additionally, if you want even more information about a certain major or certificate that is not provided by ProgressReport, such as the department heads, or general department information, all you have to do is click on the More Information About this Program button that appears when you select a major or certificate, and you will be taken to the University's Course Registrar Page.





Suggesting Courses

As mentioned earlier, after you click the SAVE button on My Board, underneath My Board will appear a list of recommended courses based on your multiple interests. For example, if you plan on majoring in Computer Science, and also plan on fulfilling a certificate in Technology and Society, Planets and Life, and/or Quantitative and Computational Biology, the after pinning these areas of interest onto My Board and clicking SAVE, you would see:

Courses we would recommend for you based on your interests..
COS323 fulfills... COS PAL
CHM202 fulfills... CHM PAL QCB
COS424 fulfills... COS TES
COS432 fulfills... TES
EGR495 fulfills... TES

This feature is meant to make the course selection process easier for students who are looking to fulfill multiple requirements with a single course.

Tracking Requirements

Another useful feature that we provide for students is to have a better idea of what requirements they are actually fulfilling with the classes that they take. When you first click on a major, a drop down list will expand to show you what tracks are within your major, as well as what classes in those tracks you have taken.

CHM *Major GPA: 3.67* Warning - Double Counting REMOVE

More Information about this Program

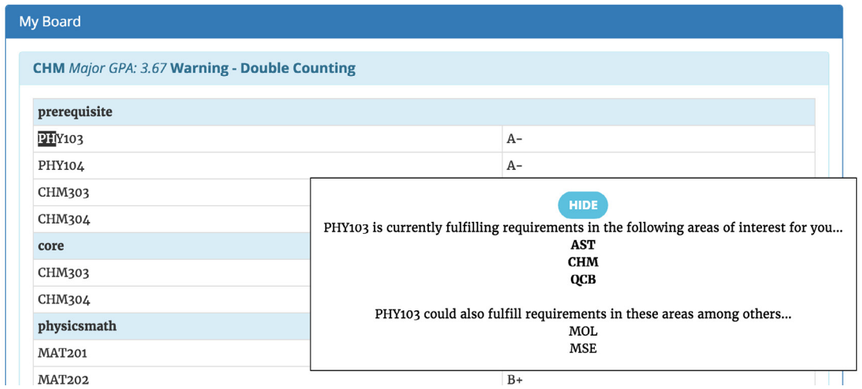
physicsmath *2 of 2 Completed*	
MAT201	B+ ✖
MAT202	B+ ✖
core *2 of 7 Completed*	
CHM303	A ✖
CHM304	None ✖
prerequisite *4 of 6 Completed*	
PHY103	A- ✖
PHY104	A- ✖
CHM303	A ✖
CHM304	None ✖

For example, the student above has completed the physics and math prerequisites to be a chemistry major, but still has 5 core classes to complete and 2 prerequisite classes.

Furthermore, if you click on any of the class names that you have taken, a pop-up will appear letting you know what requirements in your

Tracking Requirements (Cont.)

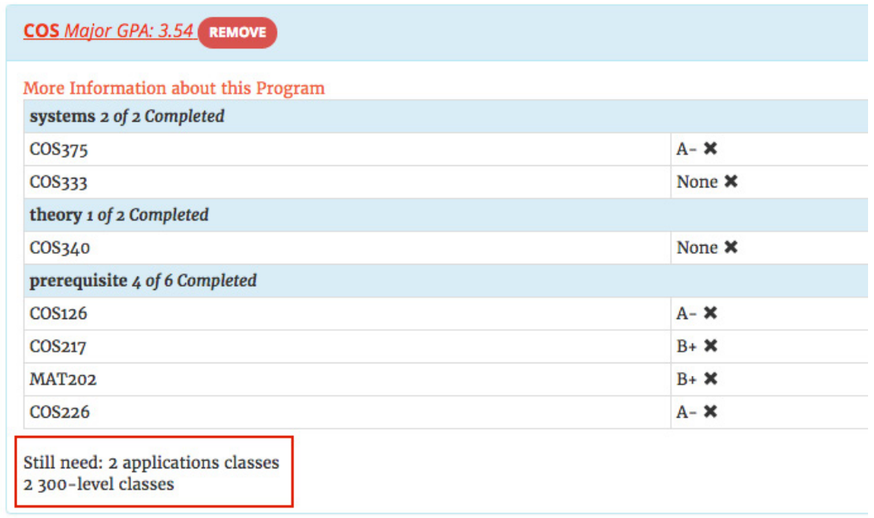
pinned areas of interest, not just the specific that youare look at in the moment, that the course is fulfilling.



The above student clicked on the class PHY 103, and can see that he or she is fulfilling requirements in AST, CHM and QCB programs, which he or she indicated interest in by pinning to their board. We also tell you a few other certificates/majors, besides those listed on your Board that this course can be used to fulfill. Hopefully, if you try this feature on a course you really enjoyed, you will get some ideas about majors and certificates you may want to pursue.

Tracking Requirements (Cont.)

We also keep track of how many other classes you still need to fulfill in the various department tracks to ensure that you have fulfilled all the requirements.



Additionally, you will notice that we provide a warning on majors where a single class is counting towards two requirements. The "Warning - Double Counting" warning means you might be lacking a course if double counting is not actually allowed in that major or certificate.

Tracking Requirements (Cont.)

ECO *Major GPA: 4.00* Warning - Double Counting PIN

More Information about this Program

departmental *1 Completed*	
ECO310	A+ ✖
core *1 Completed*	
ECO310	A+ ✖
cognate *1 Completed*	
ORF309	None ✖

For example, the student above is using both ECO 310 as a departmental and a core requirement, but it is possible that these two different requirements need to be fulfilled by separate classes, rather than a single one that is on both tracks. You can manually distribute classes that are double counted by clicking on the X to remove a double counted class from a track.

The "Warning - Overfulfilled" means that we are overcounting courses in a track and you should delete some manually or redistribute these courses to other tracks.

Tracking Requirements (Cont.)

PSY *Major GPA: 3.90* Warning - Overfulfilled PIN

More Information about this Program

200requirement *1 of 2 Completed*	
PSY254	A ✖
generalfield *2 of 1 Completed*	
COS126	A- ✖
MOL214	A ✖

In this case, a student has over fulfilled the General Field requirement for a PSY major, so she or he should select which class they want to be counted as their general field course, and delete the other to prevent this error.

We can also track how many PDFs you have used already, which can be useful to students who aren't sure if they have enough PDFs to PDF a class they are currently taking, or to let rising upperclassmen who have not used their PDFs yet to be aware of such an option.

BSE | Computer Science

You have used 0 of your 4 PDFs.

Developers Guide

For our users that are interested in continuing our work and maintaining or improving this project, this section is for you! We'll be taking you through a brief introduction to our system in this part of the guide.

Back-End Systems (Relevant Code: models.py)

As we are dealing with large amounts of data, consisting both of user data, as well as the requirements for the majors and certificates on campus, our backend consists of a postgreSQL database with five tables.

We have one table for user data, one for major requirements, one for certificate requirements, one for all courses, and one that connects majors to their required courses.

The structure of these four tables can be seen on the next page.

Database Table Structures

public.users

netid (PRIMARY KEY)	text
degree	text
major	text
interested_major	text[]
interested_certificates	text[]
courses	text[]
num_pdfs	integer
name	text
fulfilled	text[]
fulfilledcerts	text[]

public.majors

name (PRIMARY KEY)	text
track_reqs	text[]
num_pdfs_allowed	integer
total_num_courses needed	integer

public.certificates

name (PRIMARY KEY)	text
track_reqs	text[]
num_pdfs_allowed	integer
total_num_courses needed	integer

public.courses

name (PRIMARY KEY)	text
major_area	text[]
certificate_area	text[]
distribution_area	character varying (3)

public.majors_to_courses

name	text
coursetrack	text[]

Back-End Systems (Cont.)

In order to get data into these databases, we wrote and used two python scripts that parsed information from a text file into the associated database schema and input into our databases. This was a good decision so that we had a text file based backup should our database crash. We had data files for each major and certificate, in the format of:

COURSE, m (if major) or c (if certificate), MAJOR/CERTIFICATE NAME, TRACK REQUIREMENT COURSE FULFILLS

An example of a text file that contained the data for a COS major that we loaded into the database is shown:
For those requirements which were able to be

cos | parse_xml.py | controller.py

```
1  COS126,m,COS,prerequisite
2  COS217,m,COS,prerequisite
3  COS226,m,COS,prerequisite
4  COS340,m,COS,theory
5  COS423,m,COS,theory
6  COS433,m,COS,theory
7  COS441,m,COS,theory
8  COS445,m,COS,theory
9  COS451,m,COS,theory
10 COS487,m,COS,theory
11 COS488,m,COS,theory
12 COS510,m,COS,theory
13 COS306,m,COS,systems
14 COS318,m,COS,systems
```


Back-End Systems(Cont.)

fulfilled by say any 300-level class, we listed this in the database as something like DEP3 (where DEP is the department name) then used psql pattern matching to handle these special cases.

We would recommend for those implementing the system to attempt to write a webscraper to grab the course, major, track data. Unfortunately, at Princeton the major requirements pages are set up so differently from each other as to make a standardized web scraper ineffective, which is why manual collection was necessary.

All database connections and actions were handled with the psycopg2 adapter for Python.

We store not only the user's courses in the database but also their progress in different majors and certificates as calculated by us. This allows us to not only reduce computational cost by only having to rerun the progress calculation algorithm when a student re-uploads a new transcript (which is much more infrequent than the number of times a student will sign into the app), but this also allows the user to add progress to specific majors and certificates without changing their course list or affecting their prog-

Back-End Systems (Cont.)

ress in other majors in our database. Essentially, we wanted to give the user as much control and flexibility with keeping track of their progress as possible in consideration of the fact that they may be more familiar with a major than we are. For example, if the user for some reason just wants to add a course (or perhaps Independent Work progress) to one major like Computer Science without it potentially showing up in other majors such as Astrophysical Sciences, we give them the option to add to individual majors, updating their "fulfilled" or "fulfilledcerts" columns in the users table without changing their "courses" column.

We set up two inverted tables, one from majors to arrays of courses and one from courses to arrays of majors; the former was essential to our course suggestion algorithm while the latter served all other functionality such as showing what fields a course is relevant to. Finally, our majors and certificates tables kept track of the requirements needed to fulfill a major or certificate (in terms of number of courses per track, etc.), which was useful both in weighting courses for the course suggestion algorithm and reporting to the user what they still need to fulfill.

Transcript Parsing (Revelant code: controller.py - functions: parse_transcript, parse_course)

We use the Python package pdfquery to help in performing transcript parsing. In writing our commands to pdfquery, we harnessed the standard markers "GRD", "SPF", and "PDF" as our anchors to extract each individual course. Although courses appearing on Princeton transcripts differ in many respects, they all always have one and only one of the above markers on the same line as the full course information.

We used regular expressions to parse each course information line into pieces including department and number, grade, and distribution requirement. Since the registrar has slightly modified the order of course information on transcripts in the past few years, we have tested and taken special care to make our transcript parser compatible with all Princeton transcripts from at least (hopefully) the past three years. Of course, implementing this for another university would require careful attention to how their transcripts are set up.

Due to the fact that Princeton internal transcripts are initially encrypted, we had to ask the user (and provide them with instructions) to easily unencrypt the transcript before passing it to us. In the case that they did not unencrypt first or mistakenly uploaded another type of file (not their transcript), we wrote special error handling to alert the user to their mistake.

Progress Structure in Midtier (Relevant Code: Controller.py)

We structure a student's progress in the middle tier via a dictionary indexed by majors/certificates' names. Each major and certificate is itself a dictionary containing the major GPA and the courses completed in each track. We wrote a simple algorithm to sort majors by number of courses completed in the major and by GPA.

Course Suggestion Algorithm
(Relevant Code: model.py - function: suggestcourses)

1) Pool all courses that fulfill some requirement in at least one of the student's interested majors or certificates
2) Throw out courses that the student has already taken
3) Weight courses in the tracks which the student has NOT fulfilled higher than courses in the tracks which the student has COMPLETED, relative to the amount of courses the student still needs to take in that track
4) Weight courses by how many tracks they fulfill
5) Weight courses that fulfill requirements across MULTIPLE majors higher
6) Weight courses which are prerequisites higher
7) Suggest highest weighted courses (max of 5)

Front-end Overview (Relevant Code: views.py, templates folder, static folder)

Our front-end was built using Bootstrap templates. We used Jinja2 to control the information shown to the user, passing the actual majors and their course information as well as details such as which majors and certificates needed double counting warnings in a dictionary to the templates.

Although we did not achieve a fully responsive design due to time constraints, we worked to try to make our front-end as responsive and dynamic as possible. We harnessed jQuery and AJAX to this end, aiming to show changes to a student's progress as soon as they were made while reloading the page as infrequently as possible.

In our implementation, we kept separate tables for every major in every column (GPA, My Board, and Completed), meaning that updating a major in every place it appeared was more difficult. In future, we would recommend (perhaps using d3 in) keeping a major's data in one place in Jinja2 where the data can be easily manipulated and just having multiple displays of it that are all updated at once.

Conclusion

Once again, thank you for having an interest in our product and we hope that you will enjoy what we have created, whether as a casual user, or as a prospective developer!

- Marisa Chow, Isabelle Ingato, Andrew Kim & Allison Chang