

# Project Progress Report

## *Extension for Analyzing Employee Satisfaction Trend*

Name: Huijun Lin

Leader / coordinator: Huijun Lin (One person group)

NetID: huijunl2

Email: huijunl2@illinois.edu

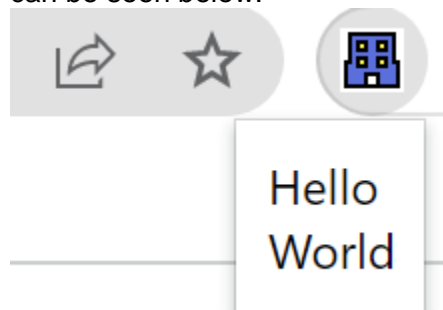
Subject Areas: Free Topics

## Summary

Task	Status
Build a "hello world" chrome extension using the provided <u>APIs</u> .	Done
Build the Crawler component	Done
Build the Analyzer component	Research
Build the Visualizer component	Not Started

## 1. Progress Made Thus Far

- Build a "hello world" chrome extension using the provided APIs (Done)
  - I followed the Chrome extension developer welcome page. I built the hello world chrome extension with only the "default\_popup" action. I learned the basic html concept and syntax following the development basic document. The end result can be seen below:



- Build the Crawler component (Done)
  - The crawler component is able to concurrently fetch a given Company's review page by page and parse the result into a list of review objects.

```
so3Language=eng
complicated so work/life balance can suffer from time to time. The difference is that it's the passion within the individuals that drives them t
work long hours and forego time off. Of course there can be times when schedule demands force the imbalance, but it's not such that 10-1
hours days become commonplace and expected regardless. We do what we do to get the job done; no one is killing themselves because
there's an order to do so.\n\nBeing able to eat to your hearts/stomachs content can be dangerous to the waistline, but that's a self-control
thing and not the company's issue.")
165. {"ratingNumber":5,"isCurrentEmployee":true,"tenure":null,"date":"Nov 9, 2021","position":"Software
Eng","location":"eer","recommend":true,"pros":"I'm still a bit new, but I have very good impressions so far. I have 10 years exp at 5+ other
companies (although no other tech companies) and this is the best by a good margin. \n\nFirst off, everyone is nice here! Yes, folks are
ambitious, but there is a shockingly small percentage of jerks here... not sure exactly how they are weeded out, but I'm grateful for it!\n\nLc
of small benefits and well-run systems give the overwhelming feeling that this company just generally has its sh*t together. Shuttles, good
food, good internal chat and groups system, not too many meetings or emails (much better than other companies!), great tech stack and
automation, etc etc. \n\nPretty good transparency and communication with leadership (e.g., weekly CEO/CTO Q&As). They are open to
feedback and change when needed, without too much unnecessary reorg-ing and pivoting. Good examples recently are moving to annual
performance reviews after years of complaints with 6-month cycle, decoupling Oculus and FB accounts, and killing facial recognition. \n\nC
is fantastic, and relatively standardized within levels. Engineers don't have different titles between levels, we're all just software engineers. TI
was weird at first, but is a really cool feature that discourages dumb hierarchies. \n\nWork life balance is totally reasonable in my experience
but might be different on other teams. \n\nMark is great, I don't really understand why folks dislike him externally.", "cons":"Some risk with r
company direction. The metaverse shift is gutsy. It seems like a reasonable move, but if it doesn't pan out, Meta might not be the same tech
giant in 10-15 years. \n\nLike any big company, it has its share of internal koolaid/propaganda (but criticism/feedback is still welcomed and
discussed openly) \n\nThe (largely overblown) negative media coverage can lead to some family and friends judging you a bit for where yo
work, if you care about that sort of thing.")
166. {"ratingNumber":5,"isCurrentEmployee":true,"tenure":3,"date":"May 24,
2020","position":"Eng","location":"eer","recommend":true,"pros":"Facebook deeply cares about its employees and has built a compelling cul
around support and growth. Career growth opportunities are plentiful. If you don't like the team you're on or don't get the support you wa
from your manger, Facebook empowers you to find new teams or projects. Facebook wants its employees to be invested in their work and t
feel connected to its larger mission. If large scale opportunities and growth are important to you, Facebook is a fantastic place to
work.", "cons":"Facebook's culture is demanding and fast paced. The greatest aspect of working at Facebook is that everyone is very motivat
and very smart. The problem with this is that they all expect the very same of you. Holding a very high bar for excellence can certainly be
demanding so it's important to make sure you're always carefully paying attention to your own personal work/life balance.")
167. {"ratingNumber":5,"isCurrentEmployee":true,"tenure":null,"date":"Nov 9, 2022","position":"Decision Science
Manager","location":"Wash","recommend":true,"pros":"- Strong growth prospects at a big company - Excellent benefits, especially health
insurance, wellness funds, and PTO - Colleagues are among the best in the world - Extremely flexible work conditions, with high degree of
autonomy.", "cons":"- Intense work experience if you want to grow - Company is run unilaterally; one man owns a commanding share of the
company and you never forget it. - Heavy-handed political stances and bias cause distrust from the consumer, and very little is done to add
it in a meaningful actions. Instead it's assumed the consumer is wrong.")
168. {"ratingNumber":5,"isCurrentEmployee":false,"tenure":null,"date":"Nov 9, 2022","position":"Data
Eng","location":"eer","recommend":false,"pros":"Open codebase, you can work on anything you want if you see value in it. Bottom up
```

## 2. Remaining Tasks

- Analyzer
  - This is still working in progress and I am now investigating which library I should rely on performing sentiment analysis.
- Visualizer
  - I have not started on this task yet. But I did a few investigations and planned to use [D3.js](https://d3js.org/) to perform visualization.

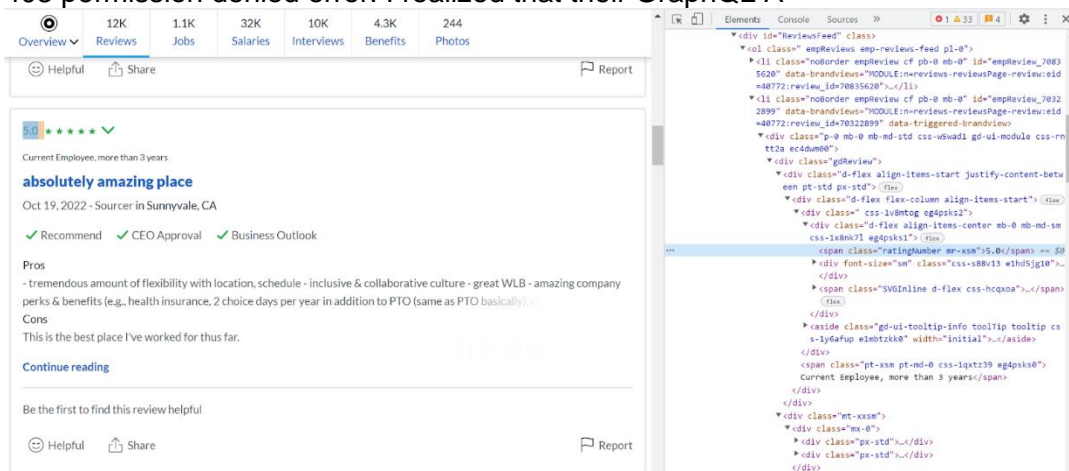
## 3. Challenges

### 3.1 Make fetch request given a glassdoor review url

By inspecting the network panel, I found that the glassdoor website uses GraphQL API (<https://www.glassdoor.com/graph>) to fetch the reviews:

```
Headers Payload Preview Response Initiator Timing Cookies
[{data: {employerDivisionReviews: null, employerReviews: {...}, featuredReviewIdForEmployer: null,...},...],...}
▼ 0: {data: {employerDivisionReviews: null, employerReviews: {...}, featuredReviewIdForEmployer: null,...}}
  ▼ data: {employerDivisionReviews: null, employerReviews: {...}, featuredReviewIdForEmployer: null,...}
    employerDivisionReviews: null
    ▼ employerReviews: {...}
      allReviewsCount: 12354
      currentPage: 3
      ► employer: {badgesOfShame: [], bestPlacesToWork: [{id: 1171, isCurrent: true, timePeriod: "2022",...}],...}
      filteredReviewsCount: 5670
      ► filteredReviewsCountByLang: [{count: 5670, isoLanguage: "eng", __typename: "ReviewsCountByLanguage"},...]
      lastReviewDateTime: "2022-11-10T11:42:57.287"
      numberOfPages: 567
      queryJobTitle: null
      queryLocation: null
      ratedReviewsCount: 3918
      ► ratings: {overallRating: 4.2, reviewCount: 3888, ceoRating: 0.77, recommendToFriendRating: 0.81,...}
    ▼ reviews: [{isLegal: true, reviewId: 70835620, reviewDateTime: "2022-11-06T18:55:23.097", ratingOverall: 5,...},...}
      ► 0: {isLegal: true, reviewId: 70835620, reviewDateTime: "2022-11-06T18:55:23.097", ratingOverall: 5,...}
      ► 1: {isLegal: true, reviewId: 70322899, reviewDateTime: "2022-10-19T20:16:10.013", ratingOverall: 5,...}
      ► 2: {isLegal: true, reviewId: 70810335, reviewDateTime: "2022-11-05T09:31:00.567", ratingOverall: 5,...}
      ► 3: {isLegal: true, reviewId: 70782945, reviewDateTime: "2022-11-04T07:12:06.427", ratingOverall: 5,...}
      ► 4: {isLegal: true, reviewId: 70789890, reviewDateTime: "2022-11-04T10:56:29.083", ratingOverall: 5,...}
      ► 5: {isLegal: true, reviewId: 70036023, reviewDateTime: "2022-10-11T16:35:58.810", ratingOverall: 5,...}
      ► 6: {isLegal: true, reviewId: 70960697, reviewDateTime: "2022-11-10T07:35:23.707", ratingOverall: 4,...}
      ► 7: {isLegal: true, reviewId: 70662527, reviewDateTime: "2022-10-31T19:29:25.577", ratingOverall: 5,...}
      ► 8: {isLegal: true, reviewId: 70767619, reviewDateTime: "2022-11-03T18:45:37.330", ratingOverall: 5,...}
      ► 9: {isLegal: true, reviewId: 6485862, reviewDateTime: "2015-05-01T00:56:22.303", ratingOverall: 5,...}
```

However, when I tried to fetch reviews using the same graph API & payload, I was getting 403 permission denied error. I realized that their GraphQL A



PI requires authorization and I have to find another way to fetch data.

By researching the glassdoor website, I found the pattern of the review page url to be:  
[https://www.glassdoor.com/Reviews/{CompanyName}-Reviews-{EncodedCompanyID}\\_P{PageNumber}.htm](https://www.glassdoor.com/Reviews/{CompanyName}-Reviews-{EncodedCompanyID}_P{PageNumber}.htm)

For example, the 2nd review page of Meta Inc. is:

[https://www.glassdoor.com/Reviews/Meta-Reviews-E40772\\_P2.htm](https://www.glassdoor.com/Reviews/Meta-Reviews-E40772_P2.htm)

I used Javascript to create the pagination url based on the current url. With the generated url, I used the [Fetch](#) API to retrieve the completed http document back, which was successful. This method takes more time to query data and requires data extraction from the html. But it was the walk around solution while the GraphQL API is not usable.



## 3.2 Parse content from the fetch response (page DOM object)

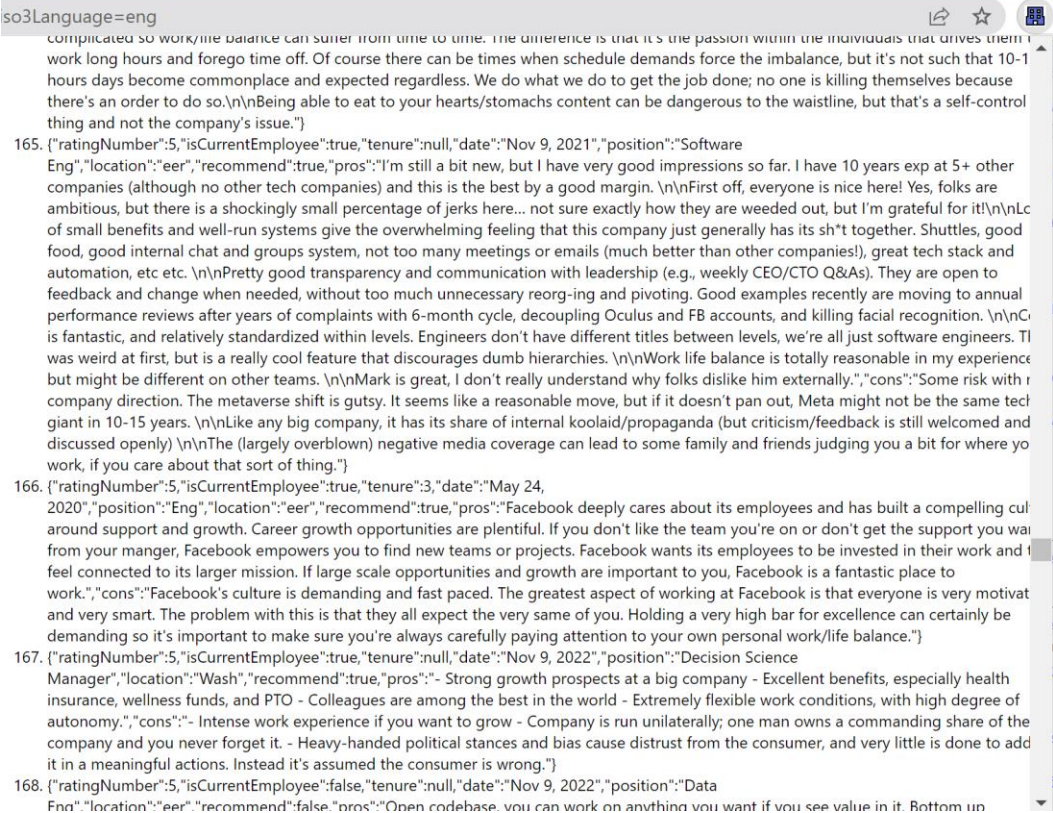
I looked into the html code and built specific logic to retrieve data I wanted (pros, cons, review scores, etc). Some of the data are difficult to locate by querySelector because they do not have unique id/class combinations. I applied the divide and conquer thinking and parsed the DOM layer by layer and eventually accurately extracted all data I required.

In addition, the reviews have different levels of details so I have to handle null values properly. The end result looked like this:

```
▼ {ratingNumber: 5, isCurrentEmployee: false, tenure: null, date: 'Nov 10, 2022', position: 'Research Intern', ...} ⓘ
  cons: "I am not sure if Meta will keep the research team in the future"
  date: "Nov 10, 2022"
  isCurrentEmployee: false
  location: "Menlo Park, CA"
  position: "Research Intern"
  pros: "research environment is very good, top researchers"
  ratingNumber: 5
  recommend: false
  tenure: null
```

## 3.3 Resolving 429 error by adding concurrency and sleep

When I tried to fetch from page 1 to N, I noticed after every few fetches, the glassdoor returns me a [429](#) error code, which means "Too Many Requests". I realized I was fetching too aggressively. I added some delay between each fetch and the error was gone. However, the fetching speed became too slow. I then added concurrency fetching to speed up the process. After a few experiments, I found that the combination of 2 concurrent requests & 1,000 milliseconds delay has a stable & relatively fast performance.



so3Language=eng

complicated so work/life balance can suffer from time to time. The difference is that it's the passion within the individuals that drives them to work long hours and forego time off. Of course there can be times when schedule demands force the imbalance, but it's not such that 10-11 hours days become commonplace and expected regardless. We do what we do to get the job done; no one is killing themselves because there's an order to do so. Being able to eat to your hearts/stomachs content can be dangerous to the waistline, but that's a self-control thing and not the company's issue.)

165. {"ratingNumber":5,"isCurrentEmployee":true,"tenure":null,"date":"Nov 9, 2021","position":"Software Engineer","location":"Menlo Park, CA","recommend":true,"pros":"I'm still a bit new, but I have very good impressions so far. I have 10 years exp at 5+ other companies (although no other tech companies) and this is the best by a good margin. First off, everyone is nice here! Yes, folks are ambitious, but there is a shockingly small percentage of jerks here... not sure exactly how they are weeded out, but I'm grateful for it! The small benefits and well-run systems give the overwhelming feeling that this company just generally has its sh\*t together. Shuttles, good food, good internal chat and groups system, not too many meetings or emails (much better than other companies!), great tech stack and automation, etc etc. Pretty good transparency and communication with leadership (e.g., weekly CEO/CTO Q&As). They are open to feedback and change when needed, without too much unnecessary reorg-ing and pivoting. Good examples recently are moving to annual performance reviews after years of complaints with 6-month cycle, decoupling Oculus and FB accounts, and killing facial recognition. The work life balance is fantastic, and relatively standardized within levels. Engineers don't have different titles between levels, we're all just software engineers. It was weird at first, but is a really cool feature that discourages dumb hierarchies. Work life balance is totally reasonable in my experience but might be different on other teams. Mark is great, I don't really understand why folks dislike him externally.", "cons":"Some risk with company direction. The metaverse shift is gutsy. It seems like a reasonable move, but if it doesn't pan out, Meta might not be the same tech giant in 10-15 years. Unlike any big company, it has its share of internal koolaid/propaganda (but criticism/feedback is still welcomed and discussed openly) The (largely overblown) negative media coverage can lead to some family and friends judging you a bit for where you work, if you care about that sort of thing."}

166. {"ratingNumber":5,"isCurrentEmployee":true,"tenure":3,"date":"May 24, 2020","position":"Engineering","location":"Menlo Park, CA","recommend":true,"pros":"Facebook deeply cares about its employees and has built a compelling culture around support and growth. Career growth opportunities are plentiful. If you don't like the team you're on or don't get the support you want from your manager, Facebook empowers you to find new teams or projects. Facebook wants its employees to be invested in their work and I feel connected to its larger mission. If large scale opportunities and growth are important to you, Facebook is a fantastic place to work.", "cons":"Facebook's culture is demanding and fast paced. The greatest aspect of working at Facebook is that everyone is very motivated and very smart. The problem with this is that they all expect the very same of you. Holding a very high bar for excellence can certainly be demanding so it's important to make sure you're always carefully paying attention to your own personal work/life balance."}

167. {"ratingNumber":5,"isCurrentEmployee":true,"tenure":null,"date":"Nov 9, 2022","position":"Decision Science Manager","location":"Menlo Park, CA","recommend":true,"pros":"- Strong growth prospects at a big company - Excellent benefits, especially health insurance, wellness funds, and PTO - Colleagues are among the best in the world - Extremely flexible work conditions, with high degree of autonomy.", "cons":"- Intense work experience if you want to grow - Company is run unilaterally; one man owns a commanding share of the company and you never forget it. - Heavy-handed political stances and bias cause distrust from the consumer, and very little is done to add it in a meaningful actions. Instead it's assumed the consumer is wrong."}

168. {"ratingNumber":5,"isCurrentEmployee":false,"tenure":null,"date":"Nov 9, 2022","position":"Data Engineer","location":"Menlo Park, CA","recommend":false,"pros":"Open codebase. you can work on anything you want if you see value in it. Bottom up