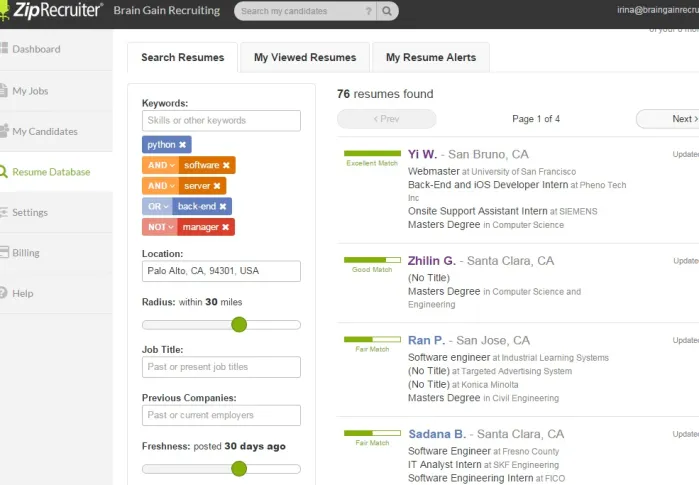
**CISC 637: Group Project - Resume Search**

**Business Problem:** Companies receive hundreds if not thousands of applications per job opening. When a recruiter or hiring manager receives that many resumes, it’s not feasible for them to carefully read each one, so typically they use an applicant tracking system (ATS) to screen applicants.

Applicant tracking systems act as a database for all potential candidates. Products such as ZipRecruiter or Indeed give recruiters the ability to manually search the database using keywords and boolean operators to find the best fit candidates.



*Example of ZipRecruiter UI*

**Technical Problem:** Create an Oracle schema that will upload a resume (word doc, pdf) and give a user the ability to text search the document.

**Username/Password Oracle:** C##\_63710\_GROUP4/C##\_63710\_GROUP4

Oracle Text Tutorials:

["Oracle Text : A Powerful Tool for Oracle Developers" by Niall Mc Phillips](https://www.youtube.com/watch?v=rfMhlKuJozI)

<https://docs.oracle.com/en/database/oracle/oracle-database/21/ccapp/understanding-oracle-text-application-development.html#GUID-CF13C01A-F5E6-4EF5-839B-C09CF0024D5E>

**Open Questions:**

* ~~Do we need a UI component (C# Blazor project)?~~
  + ~~No~~
* ~~Presentation date~~
  + ~~None~~
* Communication channel
  + Discord: <https://discord.gg/YNarN5Px>
  + Git repository - SQL scripts + data files
* Tasking out

Questions for Gibbons:

* ~~How to get files into a database. Do you need to remote in?~~
* ~~Reasonable deliverables / split~~

**Tasks**

* ~~Get the following grants from Gibbons:~~ 
  + GRANT RESOURCE, CONNECT, CTXAPP TO C##\_63710\_GROUP4~~;~~
* Grant following:
  + GRANT EXECUTE ON CTXSYS.CTX\_DDL TO C##\_63710\_GROUP4; **(Granted)**
    - Important. Need to synchronize index. Otherwise, we need to delete and recreate each time.
  + GRANT EXECUTE ON CTXSYS.CTX\_DOC TO C##\_63710\_GROUP4;
    - Allows search to return snippets of pdf. Like a couple of surrounding words.
* Optional(?):
  + GRANT EXECUTE ON CTXSYS.CTX\_CLS TO C##\_63710\_GROUP4;
  + GRANT EXECUTE ON CTXSYS.CTX\_OUTPUT TO C##\_63710\_GROUP4;
  + GRANT EXECUTE ON CTXSYS.CTX\_QUERY TO C##\_63710\_GROUP4;
  + GRANT EXECUTE ON CTXSYS.CTX\_REPORT TO C##\_63710\_GROUP4;
  + GRANT EXECUTE ON CTXSYS.CTX\_THES TO C##\_63710\_GROUP4;
  + GRANT EXECUTE ON CTXSYS.CTX\_ULEXER TO C##\_63710\_GROUP4;
* Design high-level schema - what tables do we need
  + Document, applicant, …..
* Get sample data for document table
  + 100 pdfs are enough(?)
  + Reddit ratemyresume
  + Mix pdf/docx
* Create tables
* Create sequences for tables
* Footprint triggers
* Create upload function
  + Around 21mins: 36:15 - 57:40
  + <https://udel.instructure.com/courses/1633864/external_tools/151642>

（[cisc637-010-20220412-183000.mp4](https://2358381.kaf.kaltura.com/media/t/1_xmjpgvvj/246742532)）

* Create index on document table. Video shows indexing on BLOB columns.
  + ["Oracle Text : A Powerful Tool for Oracle Developers" by Niall Mc Phillips](https://youtu.be/rfMhlKuJozI?t=2032)
* Create after insert or update trigger to synchronize index on document table
  + Section2.3.2 [2 Getting Started with Oracle Text](https://docs.oracle.com/en/database/oracle/oracle-database/21/ccapp/getting-started-with-oracle-text.html#GUID-7648841D-DE6D-441D-8C1B-5883675C236F)
* Example snippet of CONTAINS search
* Example snippet of CONTAINS with score
* Example snippet of CONTAINS with boolean operators search
  + AND
  + OR
  + NOT
* Example snippet of fuzzy search
* Example query of CONTAINS that returns applicant instead of document
* Example query with further filters on applicant, date uploaded, location, etc.
* **Requirements From Prof:**

Want a report with a description about how to do it from scratch + screen shots. (3 pages +)

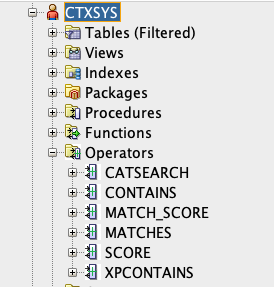
1. Build user id
2. Give permissions
3. Build the resume table and upload blobs through file upload scripts
4. CONTAINS (Xihan): Create CONTEXT index and sync index, and, or, get results by score
5. CATSEARCH - (TO DO: Tyler) DOESN’T WORK ON BLOB DATA TYPE
6. MATCH\_SCORE (TO DO: Mark) DOESN’T WORK ON BLOB DATA TYPE
7. MATCHES (mark)
8. SCORE (Xihan)
9. XPCONTAINS (Does this even exist? Can’t find any documentation on it.)
10. Index syncing behavior (Devin) - Proof of concept

1-3 are done by prof. He only needs us to mention it in the report.

5-8 are the operators in CTXSYS. He wants us to provide a description about how to use each operator and screen shot of the results.

Example from Oracl:

<https://asktom.oracle.com/pls/apex/asktom.search?tag=19c-oracle-text-indexctxsyscontext-string-search-anomaly>

[](https://asktom.oracle.com/pls/apex/asktom.search?tag=19c-oracle-text-indexctxsyscontext-string-search-anomaly)

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |