

■ Capstone Design(종합설계) 영문 계획서

| | |
|---------|-----------------------------|
| 교과목명 | Software Capstone Design |
| 팀 명 | Sprout |
| 팀 원 명 | Park Ji Soo, Kim Jeong Yoon |
| 과제명(한글) | 유전체 데이터베이스 구축 |
| 과제명(영문) | Build A Genome Database |

과제 요약(영문)

This project aims to build a genome database system related to Parkinson's disease. With the development of technology, the amount of genomic data has increased, and the construction of a database system is essential to manage such genomic data. In the genomic data obtained by researchers, there is information that is essential for security, such as patient data. The existing genome DB is open to the public, so it is not secure enough to manage and use the researcher's internal data. In this project, a genome database system is implemented to restrict who can access personal data obtained by researchers through a login system and encryption. The implemented genome database allows genome researchers to personally access their data and use it for research. Accessible data can be used by researchers to conduct their own research, such as developing new drugs.

■ Capstone Design(종합설계) 프로젝트 실행 계획서

| | | | |
|---------|-------------------------|--------|----------|
| 교과목명 | 소프트웨어 캡스톤디자인 | | |
| 팀 명 | 새싹 | 학부(과)명 | 바이오메디컬학과 |
| 과제명(한글) | 유전체 데이터베이스 구축 | | |
| 과제명(영문) | Build A Genome Database | | |
| 수행기간 | 2022년 3월 2일~6월 22일 | | |

※ 상세히 기술

1. 프로젝트 개요

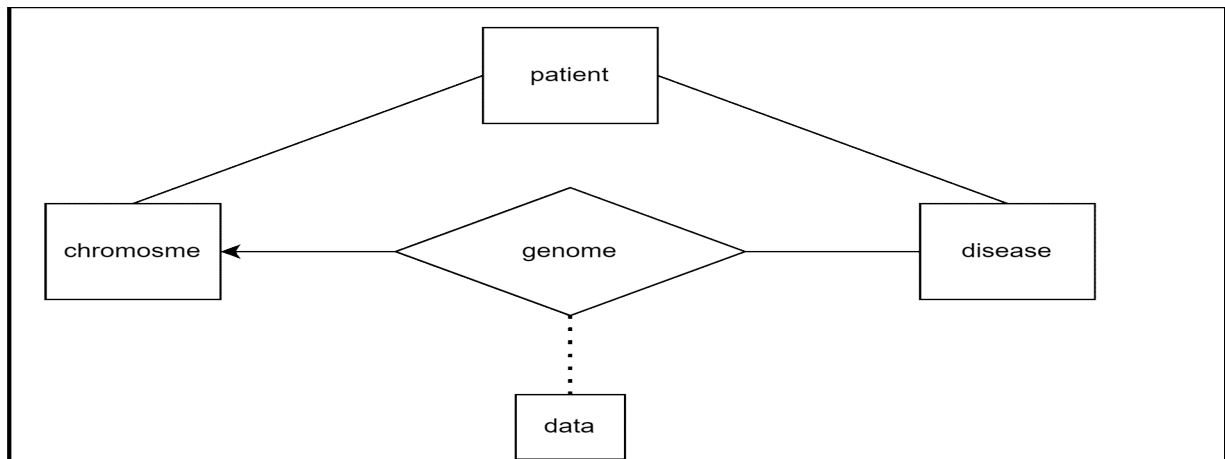
본 프로젝트에서는 파킨슨병과 연관된 유전체 데이터베이스 시스템을 구축하는 것을 목표로 한다. 기술의 발달에 따라 유전체 데이터의 양이 증가했고, 이러한 유전체 데이터의 관리를 위해 데이터베이스 시스템의 구축이 필수적이다. 연구자가 얻어낸 유전체 데이터에는 환자의 데이터 등 보안이 필수적인 정보가 존재한다. 기존의 유전체 DB는 공개되어 있어 연구자의 내부 데이터를 관리하고 사용하기에는 보안에 취약하다. 본 프로젝트에서는 로그인 시스템과 암호화를 통해 연구자가 얻어낸 개인의 데이터에 접근할 수 있는 사람을 제한할 수 있도록 유전체 데이터베이스 시스템을 구현한다. 해당 데이터베이스를 유전체 연구자들이 자신의 데이터에 개인적으로 접근해 연구에 이용할 수 있도록 한다. 접근 가능한 데이터는 연구자들이 이를 이용해 신약 개발 등 자신의 연구에 사용할 수 있다.

2. 프로젝트의 목표 및 내용(개발목표)

위 DB는 바이오 분야에서 사용될 예정이므로 해당 분야에서 요구사항을 충족시킬 수 있는 다양한 스키마를 제공해야 한다. 해당 유전체의 요구조건을 분석해 DB를 설계한다. 유전체 데이터는 인체 관련 실 데이터를 중심으로 다루고 있기 때문에 보안 또한 중요하다. 환자 ID 등 보안이 필요한 데이터를 찾고, 해당 데이터에 보안을 두어야 한다. 유전체 데이터에 접근 가능한 사람을 제한하기 위해 계정관리를 위한 로그인 DB 또한 만들어야 한다.

DB를 구현하기 위해서 MySQL을 사용한다. 해당 DB를 사용하는 사용자는 유전체 연구자일 가능성이 크다. 따라서, DB를 쓰기보다는 읽기 위주의 액세스가 요구된다. 단순한 자료의 검색이 가능하도록 DB를 만든다.

해당 유전체 DB를 웹페이지로 만든다. 웹페이지를 만들 때 DB 지식이 없더라도 쉽게 사용 가능한 인터페이스를 구현한다.



추후 ER 설계 과정 추가, 수정될 수 있음.

3. 기대효과 및 활용방안

제약회사 연구진이나 의료진 등 유전체 연구자들이 해당 유전체질환 데이터를 활용해 유전체의 정보와 단백질 발현 정도를 파악할 수 있게 한다. 이는 해당 질환에 대한 신약 개발 등에 이용할 수 있다.

3. 추진계획

1. 계획수립 및 자료조사 (3/1~3/30)

유전체 데이터베이스를 구현하기 위한 프로젝트의 큰 틀을 만들고 해당 프로젝트를 수행하기 위한 자료를 수집한다.

2. DB설계, DB구현 (MySQL이용) (3/31~중간고사 이전(4/18))

어떤 스키마를 잡을지 구상한다. 이후 MySQL을 이용하여 유전체 DB를 만든다.

만들어진 DB에서 보안이 적용되어야 할 데이터를 판단하여 암호화한다.

만들어진 DB에서 간단한 자료의 검색이 가능하게 한다.

DB를 만들 때 해당 DB를 공개할지 내부에서 사용할지 여부를 판단한다. DB를 만들 때 계정관리 DB도 포함되게 한다.

해당 유전체 DB를 웹페이지로 만든다.

3. 데이터 적용, DB보완 (4/23~5/15)

실 데이터를 적용하고, 만들어진 DB의 작동이 원활한지 파악한다.

문제가 생기는 부분이 발생하면, 해당 문제를 해결하기 위해 방법을 찾고 DB를 보완한다.

4. 캡스톤 발표 준비(5/16~)

한 학기 동안 진행한 프로젝트를 발표할 수 있도록 프로젝트에 대한 내용을 파악한 후, 해당 프로젝트를 다른 팀들 앞에서 발표하기 위한 캡스톤디자인 발표를 준비한다.

발표를 위한 자료 수집과, ppt 제작, 대본 제작, 발표 연습 등을 진행한다.

| 세부내용 | 수행기간(월) | | | | | | | | | | | | | | | | 비고 |
|--------------------|---------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|
| | 3 | | | | 4 | | | | 5 | | | | 6 | | | | |
| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | |
| 1. 계획수립 및 자료조사 | | | | | | | | | | | | | | | | | |
| 2. DB 구현(MySQL 사용) | | | | | | | | | | | | | | | | | |
| 3. 데이터 적용, DB보완 | | | | | | | | | | | | | | | | | |
| 4. 캡스톤 발표 준비 | | | | | | | | | | | | | | | | | |

유전체 데이터베이스와 관련 정보를 수집하기 위해 해당 논문이나 도서를 찾아서 참고한다.
코드 구현 등 기술 부분에서 문제가 생기게 되면 구글링과 DB관련 책을 통해 정보를 수집한다.
구글링을 통한 정보 수집에도 문제 해결이 생기지 않는 경우 DB연구실의 조교님이나 교수님,
또는 협조 업체에 문제 해결을 위한 자문을 요청한다.

4. git주소

<https://github.com/sprout-hallym2022>

5. 오픈소스사용내역

NginX를 이용한 웹서버 구축