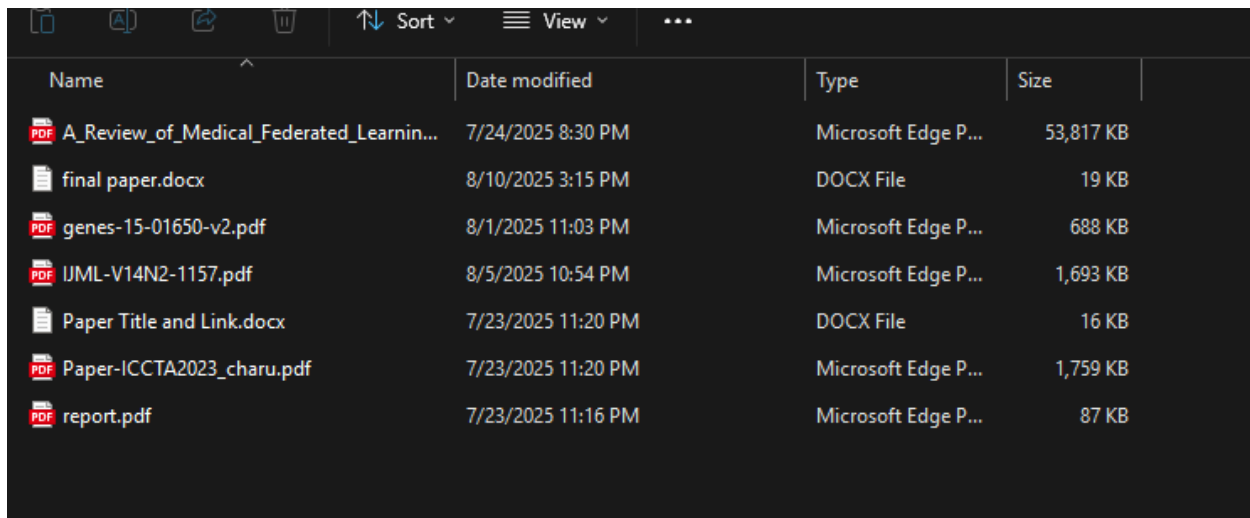









Apart from learning about Parallel Computing foundation from Distributed Computing system, I first tried connecting it with my area of interest that is Genomic Data Science. Genomic Data Science is a vast area with a focus on next generation sequencing for personalized drug discovery and health issues. My focus was on Cancer genomics. However, to fulfill my abstract I studied some articles on federated learning and genomics with some learning videos from YouTube. Upon my realization that within short time this research can't be completed, I shifted my focus on NLP, LLM hallucination within federated learning field. Here is short summary of what I learned throughout my journey of finalizing the final project as these learnings are as it was a good start for the base of my master's thesis,



Name	Date modified	Type	Size
 A_Review_of_Medical_Federated_Learnin...	7/24/2025 8:30 PM	Microsoft Edge P...	53,817 KB
 final paper.docx	8/10/2025 3:15 PM	DOCX File	19 KB
 genes-15-01650-v2.pdf	8/1/2025 11:03 PM	Microsoft Edge P...	688 KB
 IJML-V14N2-1157.pdf	8/5/2025 10:54 PM	Microsoft Edge P...	1,693 KB
 Paper Title and Link.docx	7/23/2025 11:20 PM	DOCX File	16 KB
 Paper-ICCTA2023_charu.pdf	7/23/2025 11:20 PM	Microsoft Edge P...	1,759 KB
 report.pdf	7/23/2025 11:16 PM	Microsoft Edge P...	87 KB

Here is a screenshot of learning timeline. Below is my learning journey,

Background and finding motivation:

First, I reviewed genomics and cancer's relation to it. According to WHO [1], Cancer is a leading cause of death worldwide, accounting for nearly 10 million deaths in 2020, or nearly one in six deaths. It is caused by some behavioral factors along with inheritance. However, early detection and effective treatment can cure cancer. Accordingly, many cross functional arrangements have been made to detect types of cancer. Furthermore, Digital sensing technologies are being adopted by healthcare's, enabling precise diagnosis along with data collection and processing [2]. Many approaches have been traversed by data scientists and bioinformaticians to minimize the complexity of diagnosis and personalized treatment. For instance, image segmentation using AI now gives about more than 90% accuracy in classifying cancer. [3]. Also, a good change is happening both in centralized and decentralized approaches such as deep learning and federated learning respectively. In spite of all these activities, in genomic cancer research, we are still far from adapting these tools and technologies due to many limitations. [4]. The field of cancer genomics is constantly growing and changing, fueled by the development of new

laboratory and computational technologies. However, issues such as data silos, privacy concerns, and regulatory challenges hinder large-scale data integration and collaboration. Federated Learning (FL) has emerged as a transformative solution, enabling decentralized data analysis while preserving privacy and complying with regulations such as the General Data Protection Regulation (GDPR). [5].

Dataset Exploring:

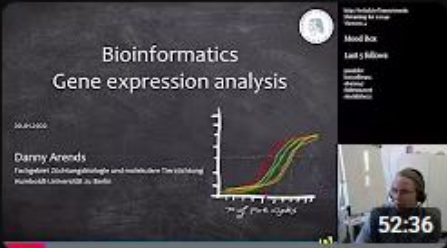
I tried exploring TCGA data and how to access those for analysis. [6]. My final goal was using a client-server environment analyzing those genomic data where my future work aimed to do parallel computing of those. Here are some videos I watched for that,

Aug 13

All views of this video removed from history

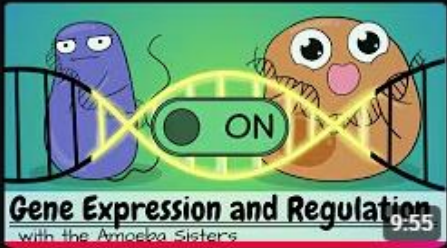
All views of this video removed from history

All views of this video removed from history



Bioinformatics Gene expression analysis
Danny Arends
52:36


Gene Expression Analysis (Bioinformatics S12E1)
Danny Arends • 3.8K views
An in-depth look at how we to measure and analyze tens of thousands of DNA probes simultaneously using RT-qPCR and micr...



Gene Expression and Regulation
with the Amoeba Sisters
9:55

Gene Expression and Regulation
Amoeba Sisters • 509K views
Join the Amoeba Sisters as they discuss gene expression and regulation in prokaryotes and eukaryotes. This video defines gene...

All views of this video removed from history



TCGA Database
11:39

Bioinformatics for beginners | Course | Download data from...
Bioinformatics Coach • 35K views
#Bioinformatics #TCGA #Genomics #Cancer This tutorial shows you how to download TCGA datasets. RNA Seq Analysis with...

References:

- [1] <https://www.who.int/news-room/fact-sheets/detail/cancer>
- [2] <https://www.sciencedirect.com/science/article/pii/S2590137025000780>
- [3] <https://dl.acm.org/doi/10.1145/3605423.3605433>
- [4] <https://www.nature.com/articles/s41576-024-00776-0#Sec7>
- [5] <https://www.mdpi.com/2073-4425/15/12/1650>
- [6] <https://portal.gdc.cancer.gov/>