

University of Essex

Department of Computing

Deciphering Big Data March

Reflective Summary

Radha Murugan

Table of Contents

Knowledge Acquired	3
Individual Contributions to the Team Activity.....	3
Experience as a Member of the Development Team	5
Impact on Professional and Personal Development	6
Conclusion.....	6
References	7

This document provides key information regarding the lectures and assignments during Units 1–12 and discusses the knowledge acquired, the experience of being a developer, how it feels to contribute to teamwork, and how it all impacts both personal and professional development.

Knowledge Acquired

Through the lectures and assignments, a variety of soft skills and technical data management skills were developed. Both the lecture sessions and the assignment on IoT (Internet of Things) played pivotal roles in fostering comprehension of the technical challenges and assessing the complexity associated with managing IoT data and addressing security issues during data exchange.

The database design assignment aimed to create a unified logical design to streamline and enhance the HR hiring process, while also conducting a critical evaluation of the pipeline. This task significantly contributed to:

- Designing the choices and **solutions** that might suit well for an organization
- Understanding the **design of the database and data management pipeline**
- Understanding the concept of normalization and building a **normalized model**
- **Data transformation** and understand the constraints and limitations associated
- **Data attributes** and data types that are relevant and critical
- **Understanding security requirements and GDPR complaint framework** applicability in exercising the rights and standards.
- Understand the emerging trends of big data analytics in modelling the large and complex data

Individual Contributions to the Team Activity

The team initially encountered a challenge in collaborating and communicating effectively, particularly due to team members residing in three different time-zones (Hong Kong, the UK, and the USA). To address this, a collaboration sheet was created to facilitate better connections and coordination among team members.

GMT	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
UK Time	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	00:00
Jeffery	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00
Danielle	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	00:00	01:00
Radha	19:00	20:00	21:00	22:00	23:00	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00

Figure 1: Contribution Allocation

To overcome the challenge of team member availability across different time-zones, multiple communication methods are being utilized to facilitate collaboration and effective communication among team members.

- A WhatsApp group was created for quick clarifications and notifications
- Team members used their personal Gmail IDs to exchange documents and references
- A Google group chat titled "DBD - Group 2" was used extensively and actively as a collaboration tool

Below is a snippet of the Google chat

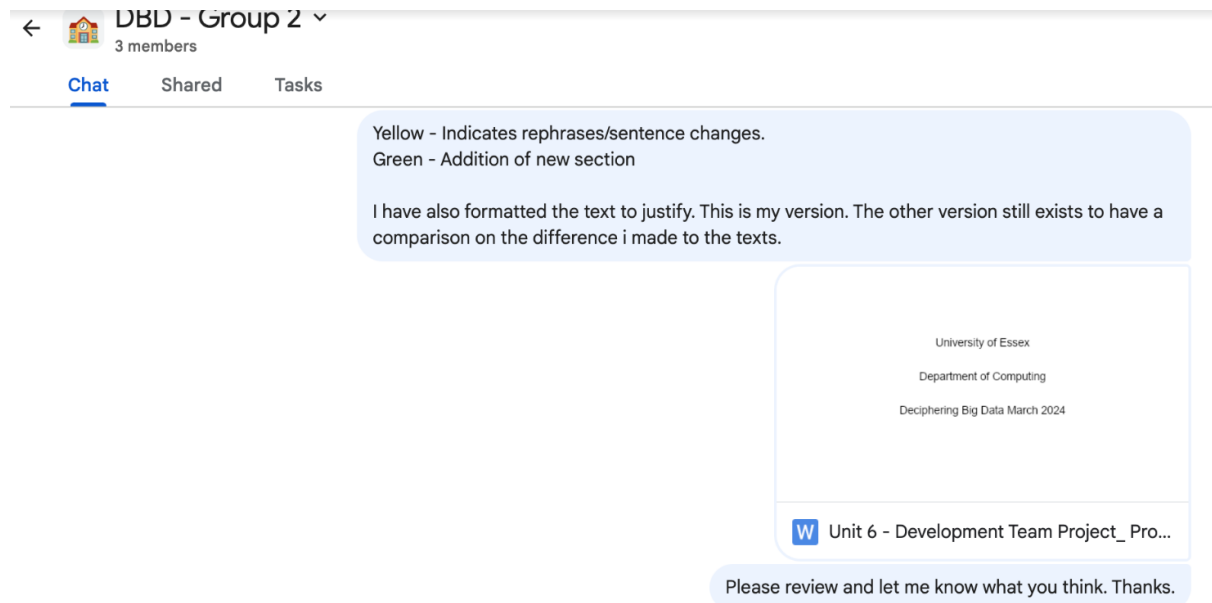


Figure 2: Google Chat Interaction - 1

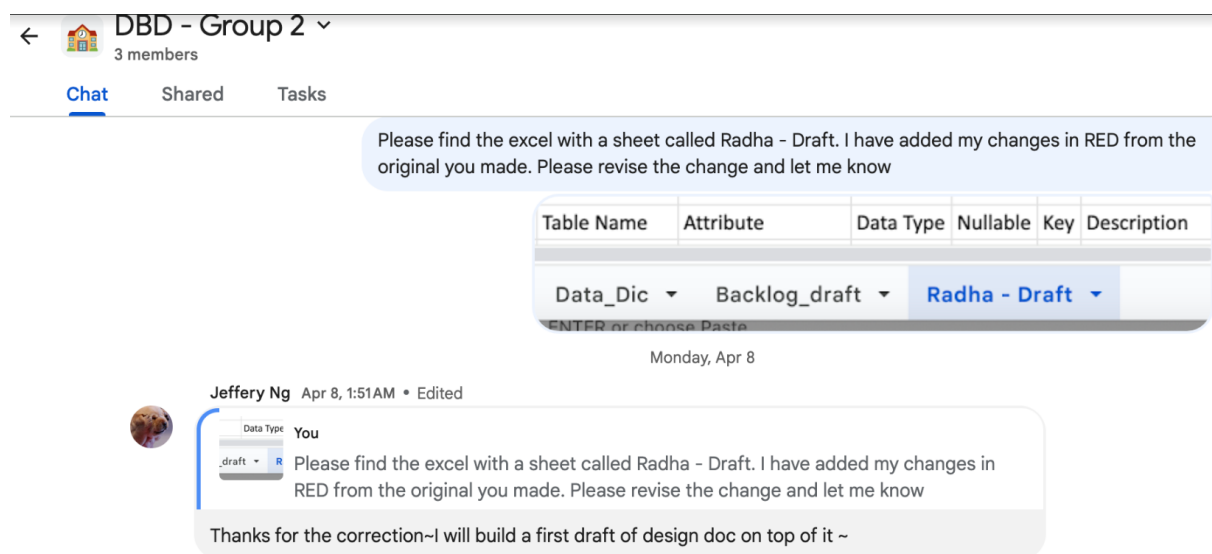


Figure 3: Google Chat Interaction - 2

The team extensively discussed the choice of organization and its objectives, ultimately selecting the HR hiring process. As part of the database design, MySQL and Oracle were considered as potential solutions (Oracle, 2021). Following a thorough evaluation of their respective strengths and weaknesses, the team decided to proceed with Oracle (Nguyen S., 2023).

The assigned tasks around analysing data attributes in detail, ensuring single-valued and unique names, and utilizing relevant data types (Romansanta J., 2020). Collaborative work was facilitated in a Google chat, where the team updated the report based on comments from an Excel document with a dedicated as tab individual contributor and then collate it together. Additionally, the introduction of the nullable option for attributes was thoroughly detailed, as depicted in Figure 4. As an individual

contributor, the quality of deliverables surpassed expectations in terms of accuracy and completeness, being acknowledged by the team.

A	B	C	D	E	F	G
Dimension	Table Name	Attribute	Data Type	Nullable	Key	Description
Employee	EMPLOYEE	EE_ID	NUMERIC	No	PK	Unique ID
		DEPART_ID	NUMERIC	No	FK	Department ID
		CAN_ID	NUMERIC	No	FK	Candidate ID
		EL_ID	NUMERIC	No	FK	Employee level ID
		EE_TYPE_ID	NUMERIC	No	FK	Employment type ID
		EE_FIRST	VARCHAR	Yes		Employee first name
		EE_MID	VARCHAR	Yes		Employee middle name
		EE_LAST	VARCHAR	Yes		Employee last name
		NICK_NAME	VARCHAR	Yes		Employee nick name
		COM_TEL	NUMERIC	Yes		Direct line
		COM_MAIL	VARCHAR	Yes		Corporate mail
		POSITION	VARCHAR	Yes		Employee position
		SALARY	NUMERIC	Yes		Employee salary
		START_DATE	DATE	No		Effective date of the record
		END_DATE	DATE	No		End date of the record
		LAST_UPDATE	DATE	No		Last update date
Dimension	Table Name	Attribute	Data Type	Nullable	Key	Description
Employee level	EE_LEVEL	EL_ID	NUMERIC	No	PK	Unique ID
		EE_LEVEL	VARCHAR	No		Employment level (e.g. entry, management,

<
>
...
Data_Dic_Revision_Jeffery
Data_Dic_Revision_Radha
Data_Dic_Revision_Combine
...

Figure 4: Data Dictionary

As part of the peer review process and it was identified that the overall topic of risk assessment (Noss, 2023) was missing from the document. Recognizing that other team members were not in a position to handle this topic, it was taken as an individual contributor additional task. The other team members appreciated for the ownership to extend support during a critical time.

The individual contribution towards team activities involved introspection and evaluation of other team members' efforts and skills, assessing how they contributed to achieving the delivery outcomes as part of the Peer Evaluation Assessment. This evaluation process helped in understanding each member's strengths and areas for improvement, as well as recognizing the collective impact of everyone's contributions towards meeting the project objectives.

It has been incredibly beneficial for fostering team cohesion and personal development.

Experience as a Member of the Development Team

Being a member of the team has been an immensely valuable journey, brimming with learning opportunities and avenues for upskilling, which have ultimately led to personal growth.

During the report delivery process, numerous challenges arose due to team members being in different time zones, making quick clarifications and coordination difficult. During the peer review process, I noticed that the document was written in conversational English, which was not suitable for producing a professional report. While rewriting the document, I encountered sentences that were unclear and difficult to understand. To avoid waiting for team members' availability, I marked these areas in different colours to revisit later, allowing me to continue rewriting and peer-reviewing the document on time. An example of this process is shown in Figure 5.

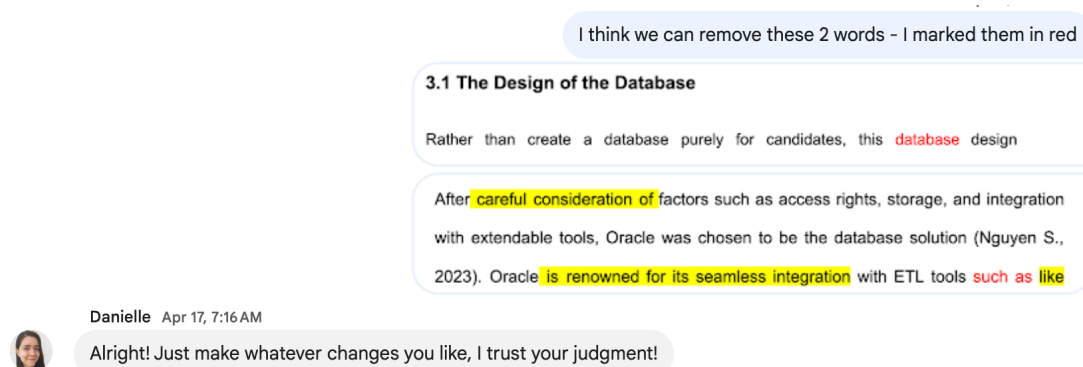


Figure 5: Example of Peer Review

Impact on Professional and Personal Development

The DB design assignment offered numerous valuable lessons and growth opportunities, both professionally and personally.

It notably elevated technical competence and deepened understanding across various domains, including data modelling, normalization techniques, DBMS, logical design, data management pipeline, and security. The experience resulted in an enhancement of problem-solving skills, particularly in handling complex and sensitive data. Designing an efficient and scalable database involved evaluating trade-offs and understanding risks to mitigate them, fostering critical thinking. Furthermore, improved effective communication and collaboration skills were gained, which are invaluable assets in practical professional settings.

Thorough analysis aids in reaching conclusions regarding data structures, identifying any missing gaps (such as the introduction of Employee Event dimensions at a later stage), understanding data relationships, and determining appropriate data storage methods to ensure the reliability of the database.

Conclusion

DB design is a dynamic field, constantly evolving with emerging technologies and methodologies, providing an avenue for continuous learning and advancement in DB technology. Through

collaboration with team members of diverse backgrounds, time zones, and skills, I learned to leverage each member's strengths, contributing to achieving the overall goal.

The individual contributions made a significant difference in achieving the overall goal. As an individual, I highly valued the opportunity to have a meaningful impact by taking on additional tasks and providing extra support to others in their tasks.

As an individual, my fellowship skills have significantly improved, particularly in areas such as data transformation, where valuable insights were offered as a team member. Additionally, my leadership skills have seen enhancement, particularly in areas like risk assessment, where effective leadership was provided to the team. The learnings from collaboration and role-played contributions have been instrumental in implementing improvements in my current developer work.

Navigating conflicts with a positive mindset and adapting to diverse situations have strengthened my resilience and contributed to both personal and professional growth. These experiences will continue to shape my journey beyond this course.

References

Oracle (2021). Available from: <https://www.oracle.com/a/ocom/docs/10-benefits-of-oracle-data-management-platform.pdf>

Nguyen S. (2023). Available from: <https://blog.dreamfactory.com/the-benefits-of-oracle-dbms-for-your-organization/#:~:text=This%20popular%20and%20powerful%20relational,a%20secure%20hybrid%20cloud%20environment.>

Noss S. (2023) Available from: Data Risk Mitigation: How To Keep Your Organization's Data Safe. Available from: <https://www.datagrail.io/blog/data-privacy/data-risk-mitigation/>