

1. Data Warehouse:

d. Data Mart:

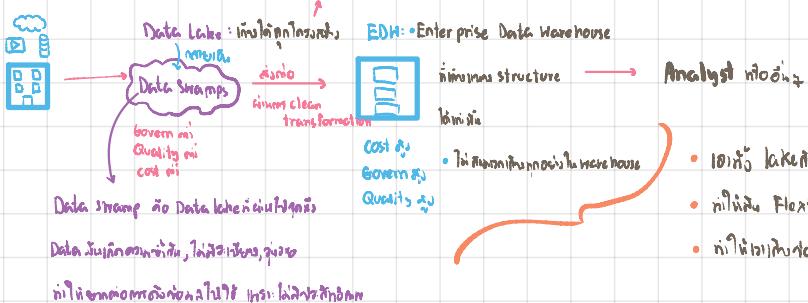
- ↳ និត្តនោយរបស់ Data Warehouse នូវ business function, purpose, និង community នូវ users
 - ↳ តុលាពាណ
 - 1. **Dependent Data mart:** ដែលនឹងពី Data Warehouse នៃ Enterprise
 - Dependent នៅលើពីរ Data warehouse នូវ Data warehouse នៃ Enterprise ដែលមានគេហទំនាក់ទំនង
 - ផ្លូវការណ៍ Clean នាម ឥតខ្ចោះពីរបាយការណ៍នៃ Data warehouse
 - មិន Data នៃគ្មានមានចំណាំពីរបាយការណ៍ទៀត
 - តួនាទី Data warehouse នៃគ្មាន នឹងជួយជាបន្ទាន់

3. Data Lake:

- ↳ **ມີຄວາມຄືຫຼຸດຈາກ**, **Architectural structure**, **semi-structure** ແລະ **unstructured**
 - ↳ **ໄດ້ກຳນົດໃຈນີ້ນັ້ນ** structure ມາ schema ນອກແຕ່ ດ້ວຍທີ່ມາ Data lake
 - ↳ **ຮັບຮັດຂໍ້ເພື່ອ data**, **ກຳນົດຂໍ້ມູນ** ພິສົດໂຄຣເກສີ່ງໄລຍະ
 - ↳ **ໃຫ້ໄດ້ການ** Data lake ທີ່ຈຳສັກຂ່າຍ, ຫຼື່ວຸ, ພິສົດໂຄຣເກສີ່ງໄລຍະ
 - ↳ **ໃຫ້ໃຈກຳນົດຂໍ້ມູນ** ລົດລົມ
 - ↳ **ໄດ້** Architecture ສິ້ນສົ່ງ **technology** ໃນ
 - ↳ **ໃຫ້ໄດ້** ດີວິວໂຫຼວດ ຕ່າງໆ ໂດຍ **ເຊັ່ນໄວ້ເວົ້າ** ໂດຍ **ມີຄວາມຄືຫຼຸດຈາກ** **Technology** ໃດ
 - ↳ **Amazon S3** **Hadoop**
Cloud Object Storage, Distributed System, Relational databases,
 - ↳ **NoSQL Repository**

Data lakehouse :

លិចនានីតិ៍ ហើយ គួរក្រោង (ពាណិជ្ជ) និងសម្រាប់អាមេរិក (ជីវិត)



- ໂຄງການ warehouse ໂຄດຖະໜີມານຸ່ງ
 - ນິຕຸມ Flexibility, cost lower, Performance, structure ທີ່ນຳເວັບໄວ້ເວັບ
 - ນິຕຸມໃຫຍ່ວັດລາງວິທີ cost lower ແລ້ວອະນຸມາ ໂຄດຖະໜີມານຸ່ງໄດ້ລາຍງານ

Considerations for choice of Data Repository:

- ↳ Use case : សម្រេចពាណិជ្ជកម្មរបស់ខ្លួនទៅជាន់តាមគម្រោងនូវការបង្កើតរឿង។
 - ↳ Data schema : និរន័យនៃការបង្កើតរឿងនូវការបង្កើតរឿង។
 - ↳ Performance requirements : លក្ខណៈការងារដែលត្រូវបានបង្កើតរឿង។
 - ↳ Data Encryption : និរន័យនៃការបង្កើតរឿងនូវការបង្កើតរឿង។
 - ↳ Volume of Data : និរន័យនៃការបង្កើតរឿងនូវការបង្កើតរឿង។
 - ↳ Storage requirements : និរន័យនៃការបង្កើតរឿងនូវការបង្កើតរឿង។
 - ↳ Organizational standards : និរន័យនៃការបង្កើតរឿងនូវការបង្កើតរឿង។
 - ↳ Capacities and access : និរន័យនៃការបង្កើតរឿងនូវការបង្កើតរឿង។

8. Compatibility: They consider how compatible the new data repository is with the existing ecosystem of programming languages, tools, and processes within the organization.

9. Security features: They evaluate the security features provided by the data repository to ensure data protection.

10. Scalability: They determine if the data repository can scale along with the organization's growth and handle increasing data volumes.

11. Skills and costs: They consider the skills available within the organization and the costs associated with different data repository solutions.

ELT, ETL, Data pipeline :

ETL: → Extract, Transformation, and Load Process

↳ និងរាយការណ៍ raw data ពីលើកទូទាត់នូវការទាំងអស់ជាន់
source និងត្រួតពិនិត្យ (Extract) និងពិនិត្យ Analyze នូវ clean, និងនូវការ
និងនិភ័យ (Transformation) ទម្រង់នូវការទាំងអស់ជាន់ និងនិងការ Load

និងនូវការ (Repository) និងនិង Automation process

↳ **Extract:** និងរាយការណ៍ source ជាប្រភេទ batch processing និងទូទាត់ខ្លួន

large chunk និងនូវការនិង stream processing និងនូវការនូវការទាំងអស់

- Stream Processing: និងនូវការនិង Transform នូវ Real-Time នូវ source

និងនូវការនូវការ Load និង Repository

↳ **Transform:** និងនូវការនូវការជាស៊ីរីស៊ីនុយោង នូវការនូវការនិងនូវការ

និងនូវការនូវការនិងនូវការនូវការនិងនូវការ Analyze

- Data Transformation: និងនូវការនូវការនិងនូវការនិងនូវការ

- Non-Data Transformation: និងនូវការនូវការនិងនូវការនិងនូវការ

↳ **Load:** និងនូវការនូវការនូវការនូវការ Load និង Repository និងនូវការនូវការ Load API

និងនូវការនូវការនូវការនូវការ Update និងនូវការនូវការនូវការនូវការនូវការ

- Full refresh: និងនូវការនូវការនូវការនូវការនូវការ reloading data នូវការនូវការ

(fresh data)

- Load Verification: និងនូវការនូវការនូវការនូវការនូវការ check នូវការនូវការនូវការ, check នូវការនូវការនូវការ

និងនូវការនូវការនូវការ

Data Integration :

↳ និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការ (ingest), transform, combine, និងនូវការនូវការ (provide) នូវការ

↳ និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

↳ និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

• នូវការ: និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

និងនូវការនូវការនូវការនូវការនូវការនូវការ

↳ **DI** Data Integration platform: និងនូវការ ETL និង Pipeline

- DI នូវការនូវការនូវការនូវការ និង Pipeline នូវការនូវការនូវការនូវការនូវការ នូវការនូវការ នូវការនូវការ នូវការនូវការ
- ETL នូវការនូវការ

Capabilities of a modern data integration platform:

- Pre-built connectors and adapters
- Open-source architecture
- Optimization for both batch processing of large-scale data and continuous data streams, or both
- Integration with Big Data sources
- Additional functionalities for data quality and governance, compliance, and security
- Portability between on-premise and different types of cloud environments

From This

Big Data :

↳ និងនូវការនូវការនូវការនូវការនូវការនូវការ Dynamic និងនូវការ, tools និងនូវការ

↳ និងនូវការនូវការនូវការនូវការនូវការនូវការ V's of Big Data

- Velocity: និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

- Volume: និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

- Variety: និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

- Veracity (准确性): និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

- Value: និងនូវការនូវការនូវការនូវការនូវការនូវការនូវការនូវការ

↳ Alternative tools: Hadoop, Apache Spark និងនូវការនូវការ

↳ Open Source to adjust Bigdata: