Title: Analysing Data Life Cycles and Modelling Data Transformation

Presenter: Claudia Roessing

Data Life Cycle

"A data life cycle provides a high level overview of the stages involved in successful management and preservation of data for use and reuse." [1]

Figure 1. DataONE

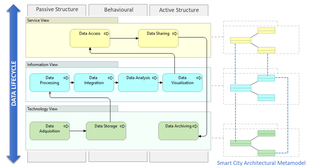


Figure 2. Smart City Data Life Cycle Sample

MOTIVATION AND OBJECTIVES

Along with the advance of technology, the appearance of big data and the introduction of new regulations like GDPR, the challenge of extracting value from and protecting data has increased significantly.

Data play a pivotal role in organizations, therefore the usage of a data life cycle is relevant, as it can assist organizations to ensure that data is collected and prepared for the intended usage and end users.

My main goal is to model a data life cycle in a Enterprise Architecture in order to assist organizations to understand the data journey.

SCOPING REVIEW

Work	No of DLCs	Criteria	Result	Disadvantage
El Arass et al. (2017)	12	- 11 phases (more phases, higher rank) - Adaptation on Big Data, Security, Green, Supervision, Management, Quality Control, Intelligence Level and Flexibility of the cycle.	Concludes that two DLCs are complimentary and recommends an union between them.	- It does not explain how the score is assigned in the second phase. - It does not say how to merge the two DLCs
Moller (2013)	8	- Phases - Features: homogeneous x heterogeneous, open x close, granularity	-	- Only states which features the DLCs have it.
Sinaeepourfard et al. (2015)	17	- 6V Challenges (Value, Volume, Variety, Velocity, Variability, Veracity).	None has achieved the goal.	- Only used the well-known DLCs from different domains, did not search in the literature

Figure 3. Analysis of Data Life Cycles

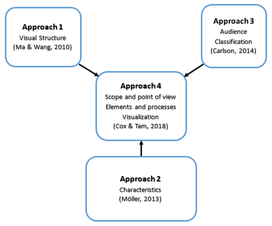


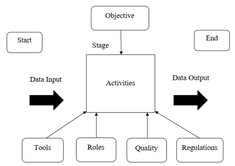
Figure 4 Analysis of Data Life Cycles Approaches

Theory

Data Life Cycle Issues:

- It is model from a high level point of view
- It is not possible to see the data transformation during the cycle
- It is necessary to have more details, therefore, it will be possible to simulate
- It needs to consider data classification, metadata, life span of data.

Data Life Cycles Requirements:



Objective - the purpose of the data life cycle, for what it is going to be used.
Stage - Represents all the steps that data needs to go through to achieve the specific outcome.
Start / End - Specifies where the data life cycle starts and where it finishes.
Data Input - The data used in every stage to be transformed.
Data Output - It represents the data which has been transformed from a previous stage and it is going to be used in the next one or it is the final output if the life cycle has reached its end.
Activities - Processes that are conducted in each stage to prepare the data for the next stage or to the final objective.
Tools - Technologies used to transform the data.
Roles - Employees responsible for each activity or stage.
Quality - It is used to know if the activities of the stage have been performed with success, if the data has achieved the goal of the stage, therefore it can proceed to the next one, otherwise it has to return.
Regulations - The data regulations that an organization has to consider in order to conduct the stages of the life cycle.

The benefits of the research:

- Modelling data life cycle in compliance with Generalized Enterprise Reference Architecture and Methodology (GERAM)
- The possibility of conducting compliance management
- Assist data managers on how to choose a data life cycle based on their needs.

Collaboration

- Contributing to the JTC 1 – WG SC40

References

[1] dataone.org/data-life-cycle
[2] Roessing , C., Helfert, M., 2019, June. The need for Mapping Data Classification Standards - Illustrated in the context of FIPS 199 and BS 10010. In Proc. of 24th EURAS Annual Standardisation Conference (pp. 397-407). K. Jacobs and P. Morone (Eds.)
[3]opengroup.org/architecture/wp/saha/TOGAF_GERAM_Mapping.htm

Ireland's European Structural and Investment Funds Programmes 2014-2020
Co-funded by the Irish Government and the European Union
European Union
European Regional Development Fund
Southern Regional Assembly