Bottom up modelling consists of 4 steps for the designing of the dimensional model[R, Kimball;R, (2013)]. They are

- Selection of the Process (Business Process or the subject of interest)
- Grain selection (Decision on to what extend data to be granularity saved)
- Decide the facts (the measurable elements or the events, or even the fact less facts)
- Confirm the dimensions

Let us equate this for the creation of an RCM Data Mart

- Selected business process is Revenues cycle management. We would like to analyze the rejection factors and to analyze the revenue.
- To analyze the data, selected grain shall be individual care item and bills.
- Fact is the encounter
- Dimension shall be based on the workflow. For example, Provider/Payer/Clinician/Patient etc.

4. Results

4.1. *Compare the models*

Inmon and Kimball, both agree on the fundamental need for a consolidated DW to facilitate a holistic data view, although the frameworks are different. Both frameworks acknowledge the importance of ETL, which will accumulate data from various operational systems and the utilization of timestamps. Mary Breslin categorize the difference between Inmon and Kimball into 3 categories based on[Breslin, (2004)]:

- The methodologies and Architecture
- The way how data is modelled
- The philosophical approach on the implementation of the DW

Let's us compare the models from a healthcare perspective now.

4.1.1. Process/Data Orientation

Healthcare industry is a very much process-oriented environment. Every provider tunes the processes in order to make the output the best. The same output could be produced in another organization through another processes. For example, have witnessed different organizations tuning the OPD appointment process in different ways to reach the same goal – "reduced waiting time. "

CIF model is not favoring process-oriented approach, it is a data driven methodology. Bottom up model is in line with this and it is completely a business process-oriented approach. Business users are involved from the very beginning for the business intelligence solution. But the issue here is the enterprise data reporting. Bottom up model is oriented towards processes than the enterprise as a whole, CIF model the other way.

4.1.2 Enterprise/Agile/Data redundancy

The CIF is a very complex model which involves the design and maintenance of the huge centralized DW. In this model, it is necessary that we decide on all the data elements in advance, which we might use once the DW is implemented. Healthcare industry is very dynamic and drawing the DW data elements beforehand is something like a mission impossible. The reason is the specific nature of the data and the changing regulatory requirements around the healthcare industry.

Kimball model is very much modular. It gives the freedom to start very small scale and build from it. Like, start with an OPD mart then extend to IP then center of excellences like rehabilitation, diabetes etc.

Inmon's enterprise view ensure that we have the single version of truth always and there is no data redundancy. In top down approach, the departmental marts are fed from the atomic EDW. That means, the Data Mart's created say OPD and IP, there is no data duplication and no chance of data integrity issue.

In Kimball's approach, the flavor of the 'single source of truth' could be lost as data is not fully integrated before it arrives the end user for analysis. Kimball model is facilitated with redundant data which can cause data update issues over time [Rangarajan, (2016)].

4.1.3 Accommodation of changing requirements.

One of the healthcare specific DW model called "Late Binding Model" by health catalyst is stressing on the need of late binding of data in the DW. Both top down and bottom up model are mapping data into a predefined model in the DW. This leads to an inability to make changes in the data model. Industries where fixed Facts and dimensions exists, this is not an issue. In healthcare business, which undergoes constant changes, early binding results in performance and capability issues.

5. Discussions and Conclusion:

We discussed the pros and cons of various DW approached with respect to healthcare. It is clear that the traditional DW and Business Intelligence approaches are not well suited for healthcare industry to adopt as it is.

If we consider industries like manufacturing, aviation, banking etc., the data analytics requirements are well defined, and top down approach or bottom up approach could be deployed as it is in their contexts. In healthcare, we lack that luxury and the industry is very dynamic and volatile.

Healthcare has two main directions, clinical is one and the other is administrative and commercial. Both these two directions to be catered equally, with respect to data analytics needs while adopting the framework.

This is leading to the requirement of a healthcare specific Data Warehouse architecture, which can cater all aspects of the healthcare industry. A combination of the existing frameworks concentrating more on the healthcare industry's specific demands is the solution. For example, the well-defined features such as RCM and financial aspects could be explored with top down model and the agile business process with a bottom up approach.

5.1. Future work of new style

Based on the conclusions made, A layered data warehouse approach with high level of integrity and holistic data view is something which healthcare demands for its long run. Authors are working on that directions to arrive into a healthcare specific DW model.

Acknowledgements

We would like to thank Mr. Ramachandran and Mr. Gopakumar for their support and motivation throughout the research work.

Conflict of Interest:

The authors confirm that there are no known conflicts of interest associated with this paper.

References

[1] Ackoff, R. . L. (1989). From Data to Wisdom. Journal of Applied Systems Analysis, 16(1), 3–9. doi: 10.5840/du2005155/629
[2] Barlow, B. S. (2014). 6 Reasons Why Healthcare Data Warehouses Fail. Retrieved from https://www.healthcatalyst.com/why-healthcare-data-warehouses-fail/
[3] Breslin, M. (2004). Data Warehousing Battle of the Giants:Comparing the Basics of the Kimball and Inmon Models. Business Intelligence Journal, 1(1), 6–20. Retrieved from http://olap.it/Articoli/Battle of the giants - comparing Kimball and Inmon.pdf
[4] Corbin, K. (2014). What Does All That Healthcare Data Really Mean? | CIO. CIO.Com. Retrieved from https://www.cio.com/article/2854599/what-does-all-that-healthcare-data-really-mean.html
[5] Dell Technologies. (2019). Solutions for Healthcare.
[6] Hamoud, A. A. K. (2018). CLINICAL DATA WAREHOUSE: A REVIEW. Iraqi Journal for Computers and Informatics, 44(2), 1–11.
[7] HIMSS. (2019). Electronic Medical Record Adoption Model | HIMSS Analytics - Middle East. HIMSS. Retrieved from https://www.himssanalytics.org/middle-east/electronic-medical-record-adoption-model
[8] Inmon., W. H. (2002). Building the data warehouse. In Wiley (3rd ed., Vol. 41, Issue 9). 605 Third Avenue, New York: Wiley Computer Publishing. Retrieved from http://portal.acm.org/citation.cfm?doid=285070.285080
[9] Inmon, W. H., & Neushloss, D. S. and G. (2008). DW 2.0. In Morgan Kaufmann (Vol. 43, Issue 6). Burlington, Massachusetts: Morgan Kaufmann. doi: 10.1016/B978-0-12-374319-0.X0001-2
[10] Lesueur, D. (2017). 5 Reasons Healthcare Data Is Unique and Difficult to Measure. Health Catalyst, 1(1), 1–6. Retrieved from https://www.healthcatalyst.com/wp-content/uploads/2014/08/5-Reasons-Healthcare-Data-Is-Unique-and-Difficult-to-Measure.pdf
[11] Ponniah, P. (2001). Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals (Vol. 6). 605 Third Avenue, New York: John Wiley & Sons.
[12] Pusic, M. (2005). Clinical decision support systems | British Columbia Medical Journal. BC Medical Journal, 46(5), 236–239. Retrieved from https://www.bcmj.org/articles/clinical-decision-support-systems
[13] R, Kimball;R, M. (2002). The Data Warehouse Toolkit - Second Edition (2nd ed.). 605 Third Avenue, New York: Wiley Computer Publishing. Retrieved from http://www.dsc.ufcg.edu.br/~sampaio/Livros/alph Kimball. The Data Warehouse Toolkit.. The Complete Guide to Dimensional Modelling (Wiley,2002)(ISBN 0471200247)(449s).pdf
[14] R, Kimball;R, M. (2013). The Data Warehouse Tool Kit - Dimensional Modelling (3rd ed.).
[15] Rangarajan, S. (2016). Data Warehouse Design – Inmon versus Kimball. Http://Tdan.Com. Retrieved from http://tdan.com/data-warehouse-design-inmon-versus-kimball/20300
[16] Warner, & Diana. (2013). IG 101: Developing Clinical Business Intelligence. Journal of AHIMA, 1(1), 1–10.
[17] Washington, U. of. (2018). Clinical Data. Data Resources in the Health Sciences. Retrieved from https://guides.lib.uw.edu/hsl/data/findclin
[18] Wasylewicz, A. T. M., & Scheepers-Hoeks, A. M. J. W. (2019). Clinical Decision Support Systems. In Fundamentals of Clinical Data Science (pp. 153–169). Cham: Springer International Publishing. doi: 10.1007/978-3-319-99713-1_11
[19] Yessad, L., & Labiod, A. (2016). Comparative study of data warehouses modeling approaches: Inmon, Kimball and Data Vault. 2016 International Conference on System Reliability and Science (ICSRS), 95–99. doi: 10.1109/ICSRS.2016.7815845