MIS362: SOCIAL MEDIA ANALYTICS AND DATA-DRIVEN INNOVATIONS.

Assignment Report.

Student Name: Nguyen Truong Huy Ta.

Student ID: 218737833.

1) **Please describe, which technological characteristics are commonly used to describe big data. Which distinct types of analytics services exist? Please, provide two examples for each type how AUDI used this type of analytics services within this case.**

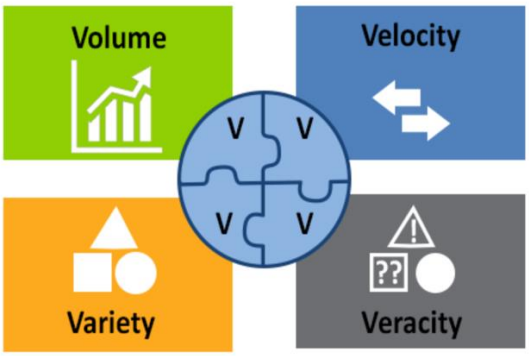
Big Data always comes with the so-called 4V which are used to describe its characteristics. Specifically, they are Volume, Variety, Velocity and Veracity.

The volume of big data is normally gigantic and requires very large databases and effective cloud servers. At Audi, data comes from different sources in large quantities such as thousands or even millions of tweets from customers about the company's car qualities, countless lists of numbers and real-time sensor data, huge sales datasets in all Audi branches in the world during recent months.

Data coming with different formats from various sources tends to present the variety of big data. It can be operational data from Audi which is structured nicely for analytics and model developments. On top of that, some unstructured data that needs to be processed and interpreted are customer tweets on social media platforms, audio data etc.

Velocity represents the speed at which data is collected, distributed, processed to produce the outcome. Specifically, it regards how quick data flows are during all the departments at Audi and whether models make predictions based on sensor data early enough to prevent product recalls.

Data trustworthiness is really important, which raises the questions whether data from a specific social media or news platforms is precise and reliable. Data that includes outliers and anomalies also needs further investigation. Those things are correlated with the term 'Variety' in Big Data.



Alvera Anto (2016) depicted Four Vs of Big Data.

Audi uses some BI tools such as Power BI and Tableau to visualize their sale performance data this month, trying to align it with their new marketing strategies. A huge amount of customer data has gone through the ETL process and is ready for customer segmentation using some clustering techniques.

Based on the predictions from ML classification model on real-time sensor data, staff would be able to identify issues with Audi cars at a paint shop. Sending advertisements to customers who are likely to purchase their products and identifying potential clients have been enabled by predictive analytics.

Diagnostic analytics may help Audi staff to detect anomalies from sensor data at Audi paint shops, enabling them to identify what is wrong with cars. The second example is that Exploratory and Diagnostic analytics is used by data scientists to identify which predictor variables make potential customers likely to purchase Audi's products.

Prescriptive analytics tell Audi which cars and how many they should produce to maximise the sale of this month but also reduce risks of products staying in stock for too long. Optimization model suggests how Audi should initiate and implement its supply chain process with a view to minimize export and import costs.

Navdeep Singh Gill (2020) depicted Predictive Machine Learning pipeline that can be applied at Audi.

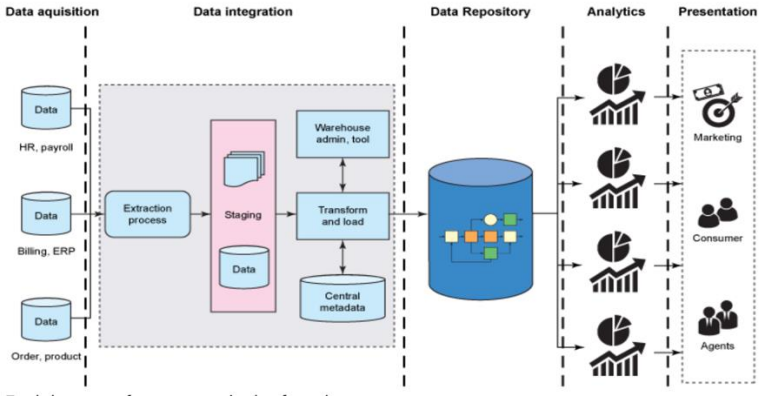
2) **Following Davenport (2006), Porter and Heppelmann (2015), and Ross et al. (2013), should AUDI follow up on a strategic path to implement big data analytics or use a step-by-step approach championed by one business department (e.g., sales and marketing) to prove and implement big data analytics? Why?**

Enterprise-level analytics, smooth data sharing and data policies are among the essential factors that contribute to a successful big data project in the long run. If all departments of Audi are familiar with how analytics are implemented and how data should be processed and stored according to data policies, it would create a really effective and transparent big data environment.

Data from all departments are processed in a standardised way and their members understand how analytics are done, facilitating data sharing among all departments at Audi and making the ETL process less time-consuming. Smooth ETL process and high-quality data thanks to standardized processing can increase big data velocity a lot, giving Audi more time to take actions after gaining insights from data analysis and predictions from Machine Learning models.

A big data project seems potential and competitive when we employ analytics throughout all enterprise departments and exchange eternal analysis with customers and suppliers (Davenport. 2006). Furthermore, analytics done at all departments as well as unified and shared across all of them not only facilitate performance self-assessment but also cross-departmental collaboration. For instance, it does not take the Supply Chain and Logistics at Audi team much time to interpret an optimisation model devised by Financial Analysts. On the other hand, Finance and cost teams also can thoroughly understand how customer segmentation and clustering analysis are done by the marketing department so that they can devise most suitable financial models for the best ROI based on customer information and preferences.

Audi is currently filled with lots department-biased projects rather than those implemented throughout all parts of the enterprise and free data and information exchange is the key to Big Data success (Dremel et.al 2018). It is understandable that Big Data tends to trivially fasten Audi decision making process and make Business Intelligence not that effective in case there is a barrier and variations in how data are shared and processed. Department-biased analytical procedures and big data approaches should be avoided and replaced by insight-generating and big data - related processes adopted at a standardized enterprise level. As a result, it would be appropriate for Audi to have a strategic big data plan covering all departments possibly using Agile Project Management methods rather than follow the approach championed by just one department.



Sandip Chowdhury (2014) mapped how Big Data infrastructure enables OLTP, OLAP and data analytics at all enterprise departments.

3) Please elaborate which of the provided Exhibits (1, 2, 3, and 4) would have been the best solution for AUDI? Why?

As discussed above, Audi should adopt a big data approach which spans over all departments of the enterprise and develop it step-by-step using Agile Management Methods. It needs to construct its own big data infrastructures with good operational databases and a data warehouse for analytical tasks and data integration regarding all Audi departments. It is also necessary that staff are trained to shift from instinctive decision making to taking action based on data (Jeanne W. Ross et.al 2009), making sure analytics and data sharing are comprehended throughout all departments.

The exhibit that is the most suitable one for the German car manufacturer is the first one.

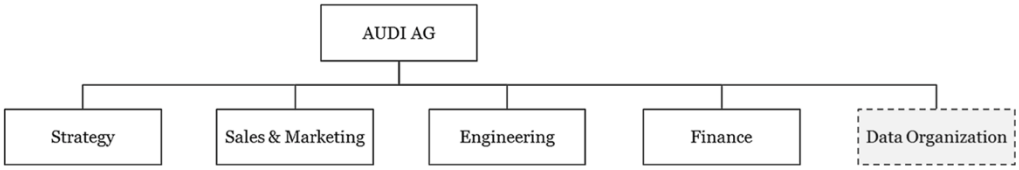


Exhibit 1 Organization chart for a unified data organization

Christian Dremel, Jochen Wulf, Annegret Maier and Walter Brenner (2018) showed the first strategy exhibit designed for Big Data at Audi.

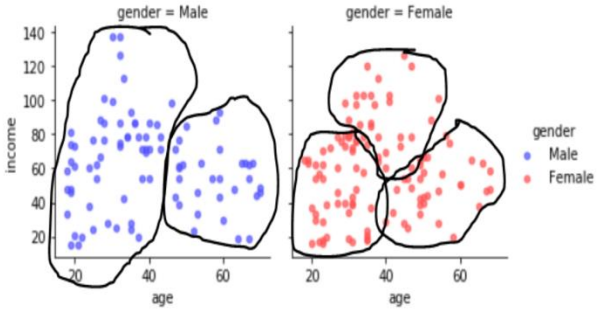
4) Which data sources of AUDI have been identified? Which are the most valuable ones and why? Please, provide at least two examples for each data source of how AUDI uses or could use this data source for innovation purposes?

The most valuable data sources at Audi are Car data and customer/web data. Car data is really essential to the car manufacturer since it includes all the information about company products which can be used for operation management and product improvement. Business is mainly about offering the best quality products to customers, which tends to be strengthened with the use of Big Data at Audi.

For instance, car data can be pulled from Audi's data warehouse to build an anomaly detection model helping staff to detect flaws and issues on vehicle engines. The combination of car data and customer information might be fed into a classification model to advise drivers for maintenance or to assess mechanical conditions with a specific degree of certainty.

Another source of data that plays an important role in Data Analytics and Business Intelligence at Audi is customer/web data. Customer/web data helps Audi staff understand customers' demands, incomes, backgrounds, etc varying on different regions in various countries so that sales and marketing departments would be able to come up with the most suitable strategy for the best revenue and ROI.

Data scientists at Audi can use k-means clustering algorithm to do customer segmentation and customer behaviour analytics, enabling them to offer the most suitable and affordable products for those living in those cities. Furthermore, predictive analytics also can help Audi prevent customer churn by telling the company which clients have the high probability of no longer using Audi services, giving the German car manufacturer enough time to take actions and retain its customers.

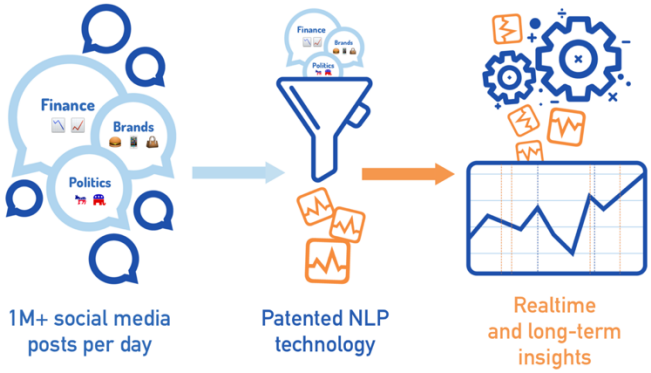


Debayan Mitre (2020) implemented Customer segmentation on Age and Income.

The least important data source among all that have been identified is third party data, including social media and socio demographic data. This kind of data does not always go with high trustworthiness - veracity, which makes it risky for Audi to use for further analytics without data filtering processes. Social media data also exists mainly as unstructured data which are tweets, audio and videos which take lots of effort as well as time to process and complete ETL process.

NLP model at Audi can be adopted using big data to crawl tweets from social media platforms for customer sentiment analysis or topic modelling, discovering people's

opinion and thoughts about its cars and services worldwide. Based on sentiment analysis, topic modelling processes and word frequencies, it is also possible for the German company to build an AI chatbot bringing better customer experience.

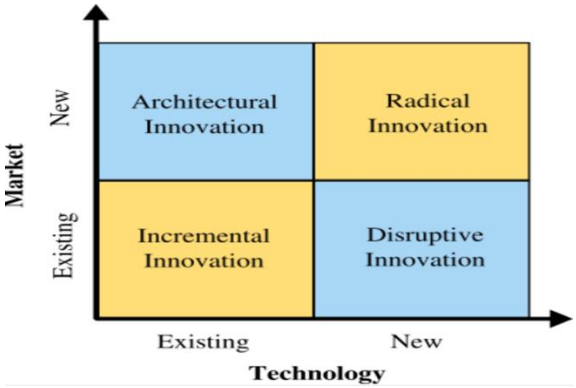


Becky Cruze (2021) - Social Media Analytics and Natural Language Processing.

5) Which innovation approach is most suitable for AUDI considering the available data sources? Why?

Audi is trying to integrate Big Data within its technological and IT system, being on the path to becoming a company using big data analytics, Machine Learning and AI at an enterprise level. Successful adoption of analytics would help Audi moving from using data merely for operational management to harnessing both OLTP and OLAP, which requires German car manufacturers to have proper ETL techniques, good data warehouse construction and insight-generating capability using Data Analytics and Machine Learning.

Instead of just storing and using data from operational databases or following decisions made by managerial people, Audi is heading towards implementing a data environment where analytics are done, insights are generated at an enterprise level and data can be smoothly integrated even if they come from different sources of various departments.



Jorge Lopez (2015) - Which innovation approach to adopt ?

The market that Audi has been and will be involved in would remain the same, which targets potential customers wanting to have decent cars and robust engines and offers them best services. The adoption of Big Data will help the company fasten decision making processes and operate in optimal ways not only based on decisions from managerial positions but also on data-centred approaches.

To conclude, it is appropriate to say that Audi is adopting disruptive innovation with its plans of Big Data Analytics, which keeps on improving after periodical iteration with the essence of Agile Project Management Method. On top of that, Incremental innovation also should be present as well since Audi needs to know how to use the existing technologies to optimize sales and offer better products and services.

References.

Thomas. H Davenport (2006), Completing on Analytics, Harvard Business Review, https://hbr.org/2006/01/competing-on-analytics

Jeanne W. Ross, Cynthia M. Beath, and Anne Quaadgras (2013), You May Not Need Big Data After All, Harvard Business Review, https://hbr.org/2013/12/you-may-not-need-big-data-after-all

Michel E. Porter and James E. Heppelmann (2015), How Smart and Connected Products Are Transforming Companies, Harvard Business Review, https://hbr.org/2015/10/how-smart-connected-products-are-transforming-companies

Christian Dremel, Jochen Wulf, Annegret Maier and Walter Brenner (2018), Understanding the value and organizational implications of big data analytics: The case of AUDI AG, Journal of Information Teaching Cases, https://doi.org/10.1057%2Fs41266-018-0036-8

Alvera Anto (2016), The 4V's of Big Data, cartographic material, retrieved 16 August 2021, https://www.zarantech.com/blog/the-4-vs-of-big-data/?utm_source=blog&utm_medium=blog-header-search-box&utm_term=

Navdeep Singh Grill (2020), Machine Learning Pipeline Deployment and Architecture, diagram, retrieved 16 August 2021, https://www.xenonstack.com/blog/machine-learning-pipeline

Sandip Chowdhury (2014), Data Warehouse Augmentation – Part 1, cartographic material, IBM, retrieved 18 August 2021, https://developer.ibm.com/tutorials/ba-augment-data-warehouse1/

Thomas. H Davenport (2006), Completing on Analytics, diagram, Harvard Business Review, retrieved 16 August 2021, https://hbr.org/2006/01/competing-on-analytics

Debayan Mitra (2020), Customer-segmentation for differentiated targeting in marketing using clustering analysis, cartographic material, retrieved 18 August 2021, https://medium.com/analytics-vidhya/customer-segmentation-for-differentiated-targeting-in-marketing-using-clustering-analysis-3ed0b883c18b

Becky Cruze (2021), iSentium Uses AI For Sentiment Analysis of Social Media, cartographic material, retrieved 18 August 2021, http://thinkapps.com/blog/development/machine-intelligence-isentium-interview/

Jorge Lopez (2015), Types of Innovation, cartographic marterial, retrieved 19 August 2021, https://techblog.constantcontact.com/software-development/types-of-innovation/