A project report on

ZS ASSOCIATES INTERNSHIP

Submitted in partial fulfilment for the award of the degree of

Bachelor of Technology

in

COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

by

Prashasti Dwivedi (171462)

To

Department of Computer Science & Engineering and Information
Technology Jaypee University of Information Technology Waknaghat,
Solan-173234, Himachal Pradesh.

Certificate of Candidate's Declaration

I hereby declare that the thesis entitled "ZS ASSOCIATES INTERNSHIP REPORT" submitted by me, for the award of the degree of Bachelor of Technology in Information Technology in the department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology Waknaghat is a record of bonafide work carried out by me under the supervision of Mr. Satyajeet Yadav, Solution Delivery Manager, ZS Associates. I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Prashasti Dwivedi

Prashasti Dwivedi
Enrolment Number: 171462
BTech (IT)

ACKNOWLEDGEMENT

EXECUTIVE SUMMARY

ZS is a global life sciences consulting firm with more than 30 years of experience, and nearly 5,000 experts in strategy, analytics and technology. Their consulting services have helped hundreds of companies worldwide solve their most critical challenges.

They bring science to bear on clients' most difficult business problems, helping them make data-driven decisions and achieve insights-fulled success. They help clients with everything from research and development to portfolio strategy, customer insights, go-to-market-strategy and sales compensation.

ZS specialises in transforming sales and marketing from an art to a science. ZS helps clients gain market share at lower cost by creating data-driven strategies that they can implement rapidly, by taking on sales and marketing operations to make them more competitive, and by helping them deliver impact where it matters.

ZS helps apply preference-driven analytics and automated tech solutions to develop and deploy a tailored customer strategy, delivering the right message at the right time, and in channels that customers prefer.

Some of the different strategies applied by ZS for customer centric marketing are:

- **Customer engagement strategy:** ZS puts customer channel and message preferences together to create data-backed engagement strategies. They use historical customer engagement data, projected marketplace changes, financial forecasting analysis and business expectations to design an optimal customer journey. They partner with the customers to deploy their plan.

- **Marketing operations**: ZS builds a direct link between customer engagement strategy and a tactical plan—ensuring that strategies and content are connected so that customer responses to one tactic will influence the next. Automation is essential to efficiently deliver messages to customers in their preferred channel, which improves engagement and ROI.

- **Customer analytics:** Their analytics solutions are built as an integrated framework to support the entire CCM process, improve overall performance and maximise revenue potential. They use:

- **Customer segmentation analysis** to deliver micro-segments that detail customer value, content, channel and offer preferences
- **Marketing mix analytics** to determine the optimal spending level by each segment
- **Customer pathway models** to measure a campaign's performance across channels for each segment over time
- **Reporting solutions** so that marketers can quickly determine if each campaign is performing as expected across each segment

- **Affinity Monitor™**: ZS created Affinity Monitor™ to provide valuable information on U.S. physician preferences for life sciences companies. Their data shows how nearly 700,000 physicians interact with e-mail, websites, personal communications and many more channels, compared with comprehensive industry benchmarks. With this rich data, their clients can tailor their promotion to each physician's channel and message preferences. The result: Increased engagement and more efficient sales and marketing investments.

- **Orchestrator Rep™:** ZS empowers sales teams with deep customer insights on channel and message preferences, information on each customer's engagement and a set of suggestions based on the latest customer insight. These are delivered to sales reps via an easy-to-use interface that works seamlessly with standard SFA tools.

- **Orchestration Engine™:** Orchestration Engine™ uses advanced data science to analyse key organizational data assets—including sales performance, customer affinity, multichannel engagement and market access—that uncover changes in customer behaviour. This analysis offers sales reps critical insights into why customer behaviour is changing and recommends the most effective course of action. Reps use these "Suggestions" to address the key issues with each customer, prioritise sales opportunities and enhance customer engagement.

- **Inside Sales:** ZS uses a data-driven approach to optimise the size, target list and call frequency of inside sales teams to yield improved customer engagement and sales growth.

LIST OF FIGURES

LIST OF ABBREVIATIONS

OE	Orchestration Engine
ETL	Extract, Transformation and Load
HCP	Health Care Professionals
CRM	Customer Relationship Management
CCM	Customer Centric Marketing
REP	Representative
DS	Data Science
AWS	Amazon Web Services
EMR	Elastic Map Reduce
NBA	Next Best Action
HDFS	Hadoop Distributed File System
PDRP	Physician Data Restriction Program
VM	Virtual Machine
BR	Business Rule
BRD	Business Requirement Document
UT	Unit Testing
SIT	System Integrated Testing
UAT	User Acceptance Testing

TABLE OF CONTENTS

1. Introduction
1.1. Objectives

Field Suggestions (FS) is an initiative to provide timely insights and suggested actions to assist reps design their customer engagement strategies and improve the effectiveness of each field interaction.

Field suggestions determines HCP level trends to provide relevant and actionable insights to reps through the Veeva platform. This automation of HCP analysis increases rep efficiency, improves the customer experience, and enables business agility.

This engagement focuses on supporting the client through a transformational Sales program to deliver high quality, business-oriented suggestions to the field through Veeva Customer Relationship Management (CRM) application iRep. The Suggestion Engine capability implementation project is designed based on the following key objectives:

1. Implement "Field Suggestions" capability based on changes in customer behaviour and cross-functional business insights to dramatically improves Sales Force effectiveness
2. Drive linkage and synergies between the Suggestion Engine project with other capability development projects within the Integrated Customer Engagement Program umbrella to fully realise the value of the program
3. Build a strong foundation of the Suggestions capability through an "Early Experience" pilot and then scale the capability across all franchises over time.

Next Best Engagement (NBE) focuses on the best sequence of actions to reach out to HCP's. It uses Machine learning/Artificial Intelligence technology. The journey designed for the HCP is based on four things:

- Customer – The target HCP
- Content – What to talk about.
- Channel – How to contact the HCP, F2F/email etc.
- Cadence – Timeline to approach the HCP.

1.2. Motivation

ZS focuses on providing clients to gain market share using data driven strategies. Field suggestion plays crucial role in interacting with the HCP (Doctor) via Rep (MR).

The analysis performed on data received based on the interaction between HCP and REP helps the client to gain access to the ground level marketing implementations and its effect. ZS's Orchestration Engine uses the data of this interaction to provide the best course of action suitable for each HCP. This project deals with a module that pushes suggestions to REP based on priority, thus helps to enhance market influence of the client product.

1.3 Background

1.3.1. Drug Life Cycle

In pharma industry releasing a drug into society is a tough task, it takes lots of testing and waiting for a drug to get into market. Lot of companies manufactures lots of drugs for same problem and out of which only few gets released into the market. FDA stated that for one disease it is said that around 5,000 to 10,000 compounds are discovered at a time, out of which only 250 of them pass the preliminary tests and go to next stage. All the preliminary testing takes up to four to four and half years. After this they go through a pre-clinical testing where the drugs are experimented. This phase takes another two to three years. After this phase only 10 to 15 drugs are sent to next phase. Now the passed 10 to 15 drugs are submitted for IND for further testing. It takes another 8 to 12 years of testing on these drugs and one or two out of them are approved by FDA and are released into our markets. A drug making company started at 10,000 compounds and ended up with releasing one or two out of them. It takes around 15 - 20 years of testing for them to release a drug.

All this process is done by the company itself with their own money. So, the amount of money they are investing in inventing a drug and going through all the testing process and releasing it into the world is in millions. Now the drug is out in the market, the testing which it has undergone is just for getting in the market, how will it reach out to its customers. There might be lots of similar drugs which are already there in market and are already been mostly used by lot of customers. How can a newly released drug counter this? In pharma industry every drug making company will have right over a drug for a time only, after this time the drug is handed over to their government and all its rights are taken from that drug making company. After the drug is handed over to the government, they release the same drug for lesser amount of money so that everyone can get benefit of that drug at lower cost only. So, the drug making companies must make as much as they can during this time only. Out of the huge amount invested for the drug making and releasing it into the market followed by the huge competition for selling their drug in the allotted time which also involves the competition with the same drugs which are released by the government. Will so many criteria under consideration a drug should earn as much as it can for its company.

1.3.2. ZS Role

The drug making companies try their level best for marketing their drugs. As mentioned, the drug making companies try to earn as much as they can from their drugs, so the normal pricing of the drugs is too high for a normal citizen to buy. So, most of these expenses are covered using insurances only.

As mentioned earlier ZS is a pharma-based consulting firm. It helps these drug making companies to get their profits back as soon as possible i.e. it helps them in increasing their market shares and help them to gain more profits. Even after doing all the marketing to the

customers it's the doctors who prescribe them to a patient. So the main aim must be the doctor who is more likely to prescribe the type of drug the drug making company has made. And it should also focus on the insurance companies who gives more amount of cut on our drug making companies product. There is lot of information which needs to be taken into consideration for making a drug's marketing successful. So ZS comes into picture here, it looks at all these constraints and will provide its clients their best marketing strategy.

1.3.3. ZS's Work

What ZS does is it helps its client the best means and the best time at which the product should be sent out to the market or the doctor. ZS help in giving the appropriate time for a representative of a client to go and meet a doctor and discuss about their drug in detail. Using this process of interaction between a HCP and a Rep, product marketing and its influence is directly taken into Consideration. Increase in product prescription benefit the Drug producing company and hence increasing their sales.

2. PROJECT DESCRIPTION AND GOALS

ZS's Suggestion Engine provides sales reps relevant and timely insights that help them develop an effective approach to customer orchestration. Suggestion Engine uses advanced data science to analyse key organizational data assets - including sales performance, customer affinity, multi-channel engagement and market access that uncover changes in customer behaviour. This analysis provides sales reps critical insights about why customer behaviour is changing and recommends the most effective course of action, referred to as a suggestion. Reps use these suggestions to address the "right" issue with each customer, prioritise sales opportunities and improve customer engagement. The following are the crucial sub-goals associated with **Suggestion Generation:**

1. **Data ingestion and transformation**: On the basis of the data received from the client, it gets ingested and then transformed for OE.
2. **Output generated as per requirement of the respective step**: After the transformation this data is as per the logic described in respective step towards suggestion generation.
3. **Configuration of suggestions in OE Centre.**
4. **Generating suggestions using OE configuration** and result obtained after transformation.
5. **Applying filters on the suggestion** as per client requirement.
6. **Pushing suggestions** in Veeva CRM: After the entire suggestion configuration in OE is completed, the final step is to Push those suggestion in Veeva CRM using which the suggestions are visible to the reps (MR)

3. TECHNICAL SPECIFICATIONS

3.1. AWS

Cloud computing is a term referred to storing and accessing data over the internet. It doesn't store any data on the hard disk of your personal computer. In cloud computing, you can access data from a remote server. Amazon web service is a platform that offers flexible, reliable, scalable, easy-to-use and cost-effective cloud computing solutions as shown in Fig 3.1. AWS is a comprehensive, easy to use computing platform offered Amazon. The platform is developed with a combination of infrastructure as a service (IaaS), platform as a service (PaaS) and packaged software as a service (SaaS) offerings.

Important AWS Services. Amazon Web Services offers a wide range of different business purpose global cloud-based products. The products include storage, databases, analytics, networking, mobile, development tools, enterprise applications, with a pay-as-you-go pricing model.

Fig 3.1 AWS Features

AWS Compute Services:

Here, are Cloud Compute Services offered by Amazon:

- **EC2(Elastic Compute Cloud)** - EC2 is a virtual machine in the cloud on which you have OS level control. You can run this cloud server whenever you want.

-
- **LightSail**-This cloud computing tool automatically deploys and manages the computer, storage, and networking capabilities required to run your applications.
- **Elastic Beanstalk-** The tool offers automated deployment and provisioning of resources like a highly scalable production website.

- **EKS (Elastic Container Service for Kubernetes)**—The tool allows you to Kubernetes on Amazon cloud environment without installation.
- **AWS Lambda**—This AWS service allows you to run functions in the cloud. The tool is a big cost saver for you as you to pay only when your functions execute.

Migration:

Migration services used to transfer data physically between your datacenter and AWS.
- **DMS (Database Migration Service)-** DMS service can be used to migrate on-site databases to AWS. It helps you to migrate from one type of database to another- for example, Oracle to MySQL.
- **SMS (Server Migration Service)-** SMS migration services allows you to migrate on-site servers to AWS easily and quickly.
- **Snowball**—Snowball is a small application which allows you to transfer terabytes of data inside and outside of AWS environment.

Storage:
- **Amazon Glacier-** It is an extremely low-cost storage service. It offers secure and fast storage for data archiving and backup.
- **Amazon Elastic Block Store (EBS)-** It provides block-level storage to use with Amazon EC2 instances. Amazon Elastic Block Store volumes are network-attached and remain independent from the life of an instance.
- **AWS Storage Gateway-** This AWS service is connecting on-premises software applications with cloud-based storage. It offers secure integration between the company's on-premises and AWS's storage infrastructure.

Database Services:
- **Amazon RDS-** This Database AWS service is easy to set up, operate, and scale a relational database in the cloud.
- **Amazon DynamoDB-** It is a fast, fully managed NoSQL database service. It is a simple service which allow cost-effective storage and retrieval of data. It also allows you to serve any level of request traffic.
- **Amazon Elastic Cache-** It is a web service which makes it easy to deploy, operate, and scale an in-memory cache in the cloud.
- **Neptune-** It is a fast, reliable and scalable graph database service.
- **Amazon RedShift-** It is Amazon's data warehousing solution which you can use to perform complex OLAP queries.

Applications of AWS services:

Amazon Web services are widely used for various computing purposes like:

- Web site hosting
- Application hosting/SaaS hosting
- Media Sharing (Image/ Video)
- Mobile and Social Applications
- Content delivery and Media Distribution
- Storage, backup, and disaster recovery
- Development and test environments
- Academic Computing
- Search Engines
- Social Networking

Data is extracted from Amazon Redshift extensively using Query language (Postgre SQL) on an SQL Workbench Platform.

AWS is very important in the field of pharma as it provides tools to create cloud-based Data Lakes for pooling, organising, and analysing real-world data for lifecycle management in pharma research and development.
 There are various AWS services which are used in the field of pharma.

3.1.1 AWS S3
Amazon Simple Storage Service (Amazon S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance. This means customers of all sizes and industries can use it to store and protect any amount of data for a range of use cases, such as websites, mobile applications, backup and restore, archive, enterprise applications, IoT devices, and big data analytics. In pharma, S3 is used for storing data which is used by the pharma companies which can be related to HCPs, sales etc. Amazon S3 provides easy-to-use management features so you can organise your data and configure finely-tuned access controls to meet your specific business, organizational, and compliance requirements. Amazon S3 is designed for 99.999999999% (11 9's) of durability, and stores data for millions of applications for companies all around the world.

3.1.2 AWS EMR
 Amazon Elastic MapReduce (EMR): EMR is a web services that provides and manages running of data processing framework such as Hadoop, Spark and Presto in an easy and effective manner. It helps to process large amount of data in a distributed environment using parallel processing.
EMR processes big data across a Hadoop cluster on EC2 and S3.

EMR has 2 nodes:

- **Master Node:** Master Node contains meta data. Meta data is a set of data that describes and gives information about other data.
- **Core Instance:** Core Instance process actual task.
- Slave node for processes
- Actual tasks

3.2. Revo Data Manager (RDM)

REVO Data Manager enables data integration and management via Amazon Web Services. Built on industry-grade Hadoop and big data technologies, REVO Data Manager empowers users at life sciences companies to access and leverage various data sources themselves. With its highly intuitive interface and defined user journey, data engineers and nontechnical users can quickly select and ingest a variety of data to conduct their analysis, all without a technical team. It supports end-to-end data services to prepare analytics-ready data sets, including data connectors, configurable DQM checks and transformation modules.

Figure 3.2 RDM Functionalities
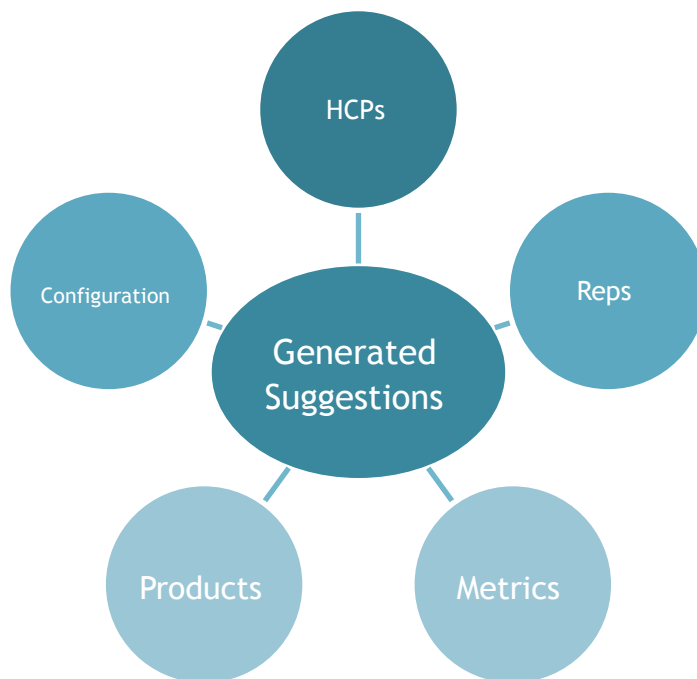
3.3. Orchestration Engine

Figure 3.3 Suggestion Generation

ZS's Orchestration Engine™ provides sales reps relevant and timely insights that help them develop an effective approach to customer orchestration.

Orchestration Engine™ uses advanced data science to analyse key organizational data assets - including sales performance, customer affinity, multi-channel engagement and market access - that uncover changes in customer behaviour. This analysis provides sales reps critical insights about why customer behaviour is changing and recommends the most effective course of action, referred to as a suggestion. Reps use these suggestions to address the "right" issue with each customer, prioritise sales opportunities and improve customer engagement.

Features:

- The prioritised suggestion list based on customer, headquarter and rep initiatives helps reps address the "right" issue and leads to more effective sales calls
- A machine-learning algorithm delivers improved suggestions over time based on adoption of suggestions, sales rep feedback and multichannel data updates
- Includes a predefined suggestion library with 60-plus insight-action pairs across relevant commercial data area.
- Standard data adapter and data ingestion framework for third party marketing vendors

- The Orchestration Engine™ administrative console contains prebuilt reports that help analyse rep adoption and correlation between rep usage customer engagement and sales
- Rapid implementation methodology and lean operations model
- Flexible cloud-based architecture allows you to extend and scale Orchestration Engine™ across business units and geographies
- Agile implementation methodology and global operating model.

3.3.1 Provide sales reps timely, relevant and actionable suggestions

Reps can use suggestions within Orchestration Engine™ to quickly refine their customer engagement strategies.
- Provide proactive and predictive customer management with early detection of opportunities and risks
- Significantly enhance rep efficiency with insights that are systematically linked to business priorities and diagnostics that help reps address the "right" issues with customers.
- Better alignment between field execution and brand strategy.
- A customised rep experience provides increased rep adoption leading to higher HCP engagement.

3.3.2 Make sense of data and deliver intuitive suggestions with statistical analysis

- Analytic techniques are used to detect changes in customer behaviour and trigger suggested actions for reps to take.
- Correlated multilevel insights and diagnostics provide a rich perspective that would be difficult to obtain manually.
- Orchestration Engine™ incorporates feedback from reps on individual HCP suggestions to improve the relevance of suggestions delivered over time.

3.3.3 Deploy quickly and efficiently with a flexible, cost-effective solution

- Flexible delivery and integration options help you reap the benefits of this solution quickly
- We offer comprehensive sales force behavioral training and change management to help enable rep and organizational adoption.

3.4. Veeva

Veeva has been Salesforce's preferred worldwide CRM provider for the pharmaceutical and biotech industry.

Built on the Salesforce Platform, Veeva CRM is core to Veeva Commercial Cloud, which brings together customer data, compliant content, and multichannel engagement for life sciences companies to deliver the experience healthcare professionals have come to expect. The integration of Veeva CRM and Salesforce Marketing Cloud allows information to be shared between the two cloud solutions so that sales and marketing teams can have a complete view of customer engagement. We use veeva to push suggestions, i.e. to show the generated results to a rep.

It has the following components:

- Suggestions - Contains all information about the suggestions sent to sales reps.
- Accounts – Contains basic information about HCPs.
- User - Contains basic information about sales reps.
- Feedback – Contains reason for dismissal of suggestion.
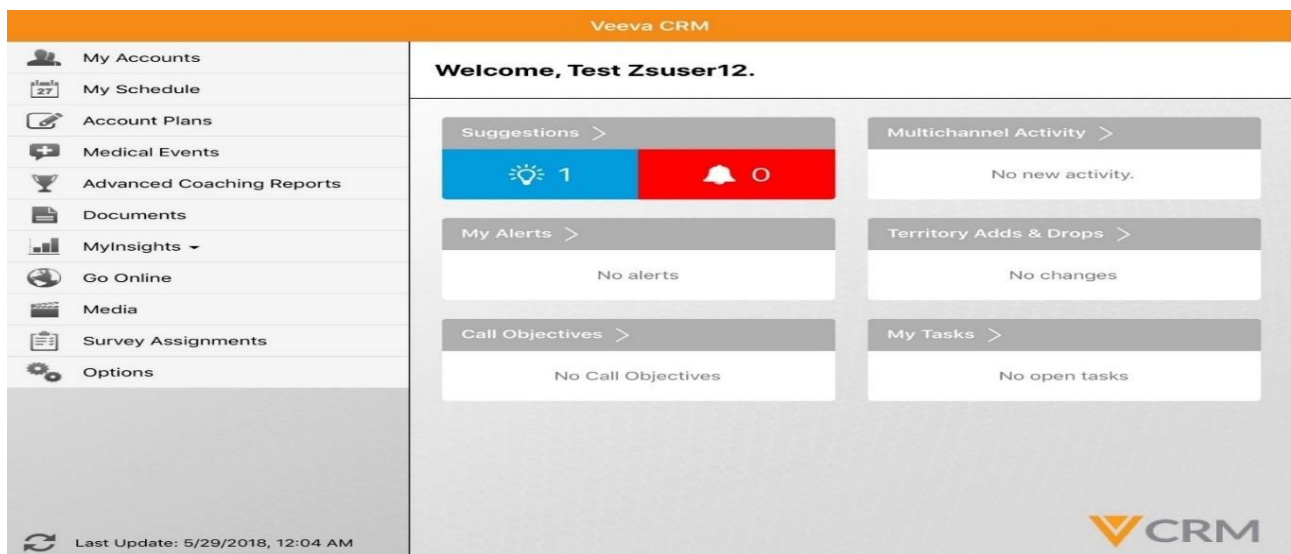- Tags – Contains product name, expiry date.

Figure 3.4 Veeva CRM

3.5. Apex Data Loader

Data Loader is a client application for the bulk import or export of data.

It is used to insert, update, delete, or export Salesforce records.

When importing and exporting data, Data Loader reads, extracts, and loads data from comma-separated values (CSV) files or from a database connection.

In this project, it is used for Export/Insert/Update/Delete suggestions/tags pushed in Veeva.
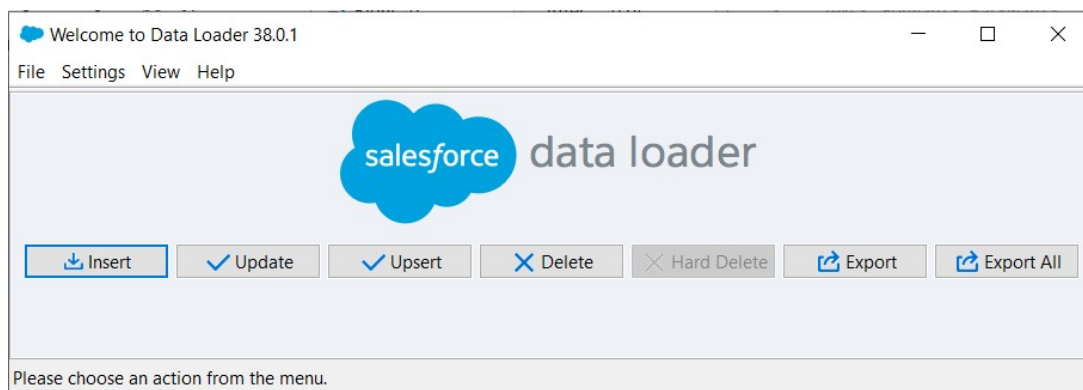

Figure 3.5 Apex Data Loader

4. DESIGN APPROACH AND DETAILS
4.1. Design Approach/Material & Methods
4.1.1 Orchestration Engine Flow

Below figure represents the High-Level Design of Orchestration Engine. Following is the sequence of this design.

- Data: Received from client in a definite format decided with client and suitable to be directly used in Orchestration Engine.
- Admin Portal/Configure Engine: Here we select the HCPs on which the conditions are to be applied. This is followed by describing the thresholds for different insights based on which the suggestion is triggered.
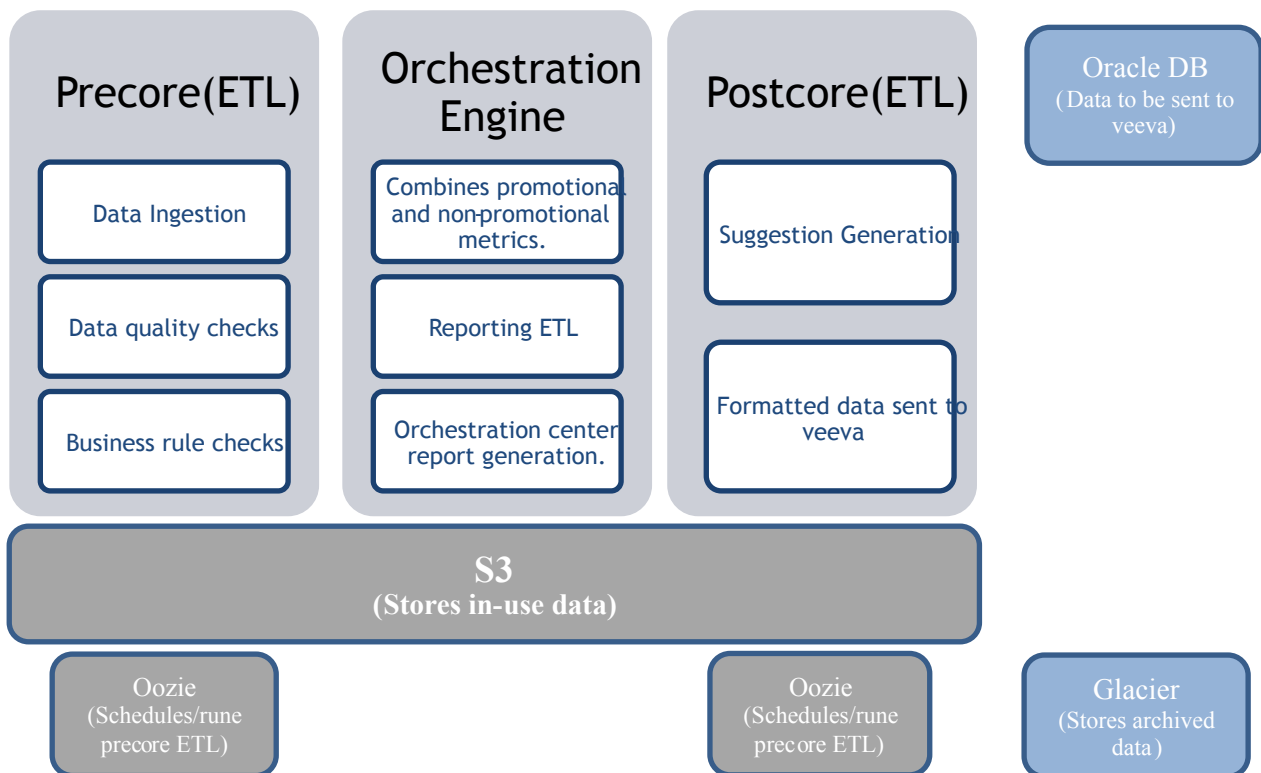
Figure 4.1 OE Overview

- OE: It uses the client data and the configure engine rules to generate suggestions for the HCP and then aligns it to single or many reps as per requirement.

- Veeva CRM: The suggestions generated are pushed to Veeva, which is a portal for Reps to check action for particular or aligned HCPs. All the feedback corresponding to the suggestion is sent through here.

4.2. Codes and Standards

- Data
- OE is constrained to obtain data in a particular manner.
- OE requires 6 input files in respective formats to be able to use effectively. o 6 Files: NPP, PP, Assignment, Rep Data, HCP Demo and HCP Segment.
- Content
- Content used in HCP-REP interaction is limited to the approved document.
- Any small change in the content of this interaction is not accepted.

⦿ Suggestion

⦿ No. of suggestions sent to a rep are strictly restricted to the approved amount.

⦿ Number of HCP reachable to Rep is strictly monitored.

4.3. Constraints, Alternatives and Tradeoffs

Data is the key item which drives the complete OE, this data is the client related data on which the client wants to perform some operations and increase his business needs. Each client might have different format of data and we cannot use different formats of data as inputs to our OE, so we made a standardized data input format which our OE requires as its input files and it uses these input files only and performs its operations. There are seven standardized input files which OE accepts, they are:

1. Personal Promotion
2. Non-Personal Promotion
3. Rep Data
4. Assignment Data
5. HCP Data
6. HCP Segment
7. Log Files: They contain the logs of all the above files.

Brief description about the above files are given below.

1. **Personal Promotion :** This file has the data about the HCP and Rep interaction i.e. data like when did a Rep go and had a talk with an HCP.

2. **Non-Personal Promotion :** This file consists of the data about HCP and metric/ variable E.g. Increase in sales of competitive drug or decrease/increase in sales of client's drug for HCP. Here sales of client's drug are a metric, increase/ decrease in the measure of that particular metric.

3. **Rep Data :** The details of the Rep who are affiliated to our client are present here. Using this file, we get to know which Rep is assigned to which region or which Rep. E.g. Name of Rep, territory, district and which franchise he is assigned to.

4. **Assignment :** This file has the data about the alignment of the rep and HCP. Here customer ID is the HCP ID.

5. **HCP Data :** As Rep Data consists of details of Rep's present with our client,

6. HCP Data has all the data about the HCP's our client would like to contact. PDRP is a flag which determines whether a rep can contact that person or not. If this flag says 'N'

then the HCP has all rights to file a case when a person visits him for any drug publicity purpose.

7. **HCP Segment :** This file consists of the data about which HCP is assigned to which Segment. Segment is a type of classification that HCP's are classified into.

8. **Log Files :** Log file keeps the log of all the files mentioned above. These data files are provided to the Orchestration Engine along with the rules which are configured from the Admin portal and the OE configures the suggestions based upon them. Whatever might be the data format of any client, we perform few steps and get the data in the form of these seven files.

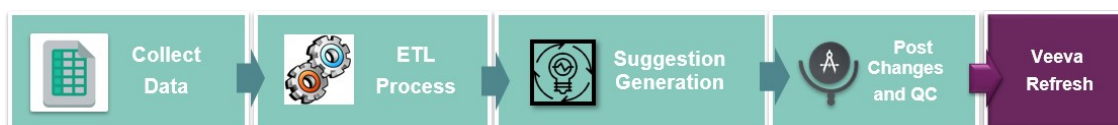# 5. PROJECT DEMONSTRATIONS

5.1   Flow in OE core Engine

Figure 5.1 OE Flow

Now we have seen what all the individual parts of the complete Orchestration Engine do, let's have a look at the process flow of the OE. Below is the stepwise description on the steps in OE. For triggering the cluster/AWS machine, an empty file is kept in the AWS core engine and it is executed, this empty file will trigger the AWS machine and the process gets started. The first step it will do is:

- **ETL Step :** Now that we know how the AWS architecture is working, let's see the flow of the OE data. First the raw data is obtained from our clients which is not present in the six input files format. We assign a task to our OE core machine to convert these data files to those six input files formal. This step is known as ETL process. We use Flow Builder for making the data transformation into these files. It is UI based component of Orchestration Engine. It has source file nodes (for ingesting source files), source table (for creating source table), temp nodes(to create temporary table/view), SQL nodes(for transformation through SQL queries), remote table(for using tables created in other flows), action nodes(for executing entire flow), salesforce nodes(for ingesting data through salesforce). It is used to read data from a file and temporarily store it in memory as a table and perform operations on that table and then store back into a table or a file. This step is known as ETL (Extract, Transform and Load) step. The existing data is extracted, transformed based on our requirement i.e. transforming into six files and the

loading into the OE. After performing the ETL step by the core engine we are left with six data input files which are then stored back on the S3 location. Now after obtaining these six files we go into the next step.

- **Data Transformation :** Data Transformation step consists of creating a pivot up combined data lake for all the OE input files. It consists data from Personal Promotion, Non- Personal Promotion, HCP Data and HCP Segments file. Data is grained at HCP level. Entire data is then passed to Prioritisation process. Reading data from six large input files, whenever required, is a huge task even for a big machine-like AWS. So, we try to merge all the data files into one by the above mentioned process.

- **Suggestion Prioritisation :** It involves generating suggestions and assigning priorities by STAR algorithm: - • Severity Timeliness Actionability Relevance .Each suggestion generated is given priority, based on the above mentioned factors. It is the most important step during the entire process. It consumes ETL output data as well as suggestion configuration on OE portal provided as json file.

- **Suggestion Suppression :** Same suggestion should not be generated every week. There might be the case when there is no significant change in the sales of drugs and would lead to suggestion generation. But we do not want same suggestions to be pushed for every run and thus suggestion generated in the previous runs should be suppressed for specific duration e.g. 1 month. Here we use the archived suggestions from previous run and search for HCP Id and Theme combination. If it is found, then suggestion in the current run is suppressed, else it is passed to the next step.

- **Limit Suggestion :** Suggestions generated are provided to the Field Reps. Reps can complete the given suggestion within some limit. If we bombard suggestion for 1 week and then send 0 suggestions next week then it does not make any sense. So we try to cap the suggestions generated according to clients requirement. Reps would only be able to get 15 / 20 / 25 suggestions based on fixed limit. Suggestions are selected based on the theme rank given to each theme. For example, Competitive drug A is more harmful for the sales of our client than competitive drug B. So, we give higher rank to competitive drug A theme and lower to competitive drug B theme and thus suggestion for higher theme ranks would be pushed capped by static limit number.

- **Veeva Ready File Generation** : Veeva is a CRM application which is used by every rep to find the appropriate suggestion for him to go and interact with the HCP. Suggestions are generated and then pushed to Veeva (An app used by the Reps to get the suggestions pushed by ZS). We have two-way API – based data connector with Veeva. We convert suggestion generated into veeva ready format and then push suggestions to Veeva app through data connector. After pushing suggestions reps are able to see those suggestions in their iPad. For every suggestion, there is linked suggestion survey which takes feedback from the reps.
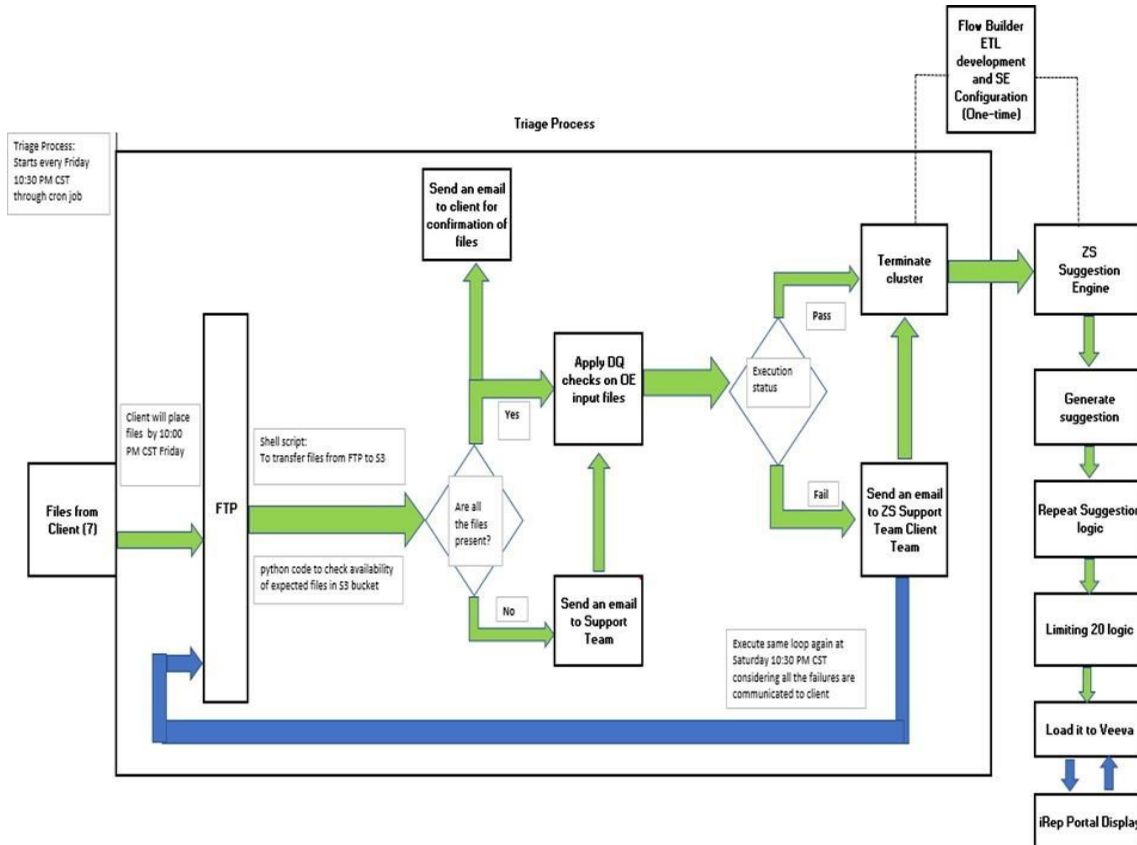

Figure 5.2 OE Suggestion Generation Flow

6. RESULT & DISCUSSIONS
6.1. Revo Data Manager (RDM)
REVO is an ETL tool used for data management.

- **Extraction and loading** are done using Databases – Oracle, Teradata, HDFS, Redshift
- **Transformation** is done using Queries – SQL, Python, Java
- **Reporting** is done using – Tableau, Micro-strategy, Spot-fire Data Ingestion and transformation is performed here. Data from the client is transformed to be suitable for Orchestration Centre. All the product step related data transformation is done here and

functioning of non-product step is defined here. It acts as a process flow for the entire run. It indicates the procedural execution of the process and resultant a successful run.

All these flows are created in the backend as a json which includes all the queries and all the flows part of this entire system. This json file is then used to execute stepwise command in EMR Cluster using the unique id for some steps. The output location is provided by the developer which help to validate the results after the run.

Components of RDM
- **Adapter:** Adapter is used to establish connection with different data sources such as S3, FTP.
- **Ingestion:** Ingestion is the process of pulling data files from data sources to HDFS.
- **DQM:** In Data Quality Management, different quality checks are applied. The types of quality checks we do is null value checks, unique checks, defaults checks.
- **BRE:** Business rule executor contains business logic to transform the data according to business requirements.
- **Export:** In this phase we push the data files back to S3 after the processing is complete.

6.2. Configuration
After getting all the output using RDM, the next stop of generating suggestions follows. For this, we have to define several rules and triggers along with its thresholds, to decide on which HCP-REP combination must be triggered and for which theme shall it trigger.
The rules which are to be applied on the six input files are configured from this admin portal. It is a dot net(.Net) based UI application which is hosted on Amazon Web Services Beanstalk web server and it uses SQL server for storing its data till R8 upgrade, but from R9 upgrade the data of configure files is stored on SQL server and data which consists of rules. The configured rules in this portal becomes another json called
'rules.json' (or similar as per need). This json comprise of all the rules that are configured on the UI. It provides theme-wise details of the rules that are applied to provide suggestion. The Orchestration Engine hits the SQL server and asks for the rules.json file which is used for configuring the output suggestions. This SQL server then asks the Oracle Config schema and returns the data to the Orchestration Engine. The data which is obtained in this process(rules.json) is used for configuring the OE. The admin portal is not only used for configuring the rules but also used to find out the stats of the suggestions and the product. For configuring the rules.json file, we need to have particular set of values which needs to be pre-configured into our admin portal so that we can select those values for the respective product. There are 36 files which are to be pre-configured before setting up the rules.

After configuring these files into our OE we can then configure our trigger conditions on which our metrics should get triggered and these conditions are imported from this OE into our chore engine as rules.json file and these conditions are applied on the input files to get suggestions.

Apart from this any suggestion is classified into two types: A **Business Rule** (BR) suggestions is type of suggestion which has trigger which are based on metrics where as a **Data Science** (DS) suggestions are the type of suggestions which are made using the comparison between the own product or its competitor. If we have any comparison type of metrics, they fall into DS suggestions and rest will be present in BR type of suggestion. Other important and crucial feature of the admin (Configuration Centre) portal are:

1. **Insight Text**

Provides an insight into the reason behind the suggestion, it gives the wording for the suggestion along with the data level stats for the suggestion trigger.

2. A**dvanced Insight**

For some cases insight text varies according to the threshold condition, due to which same type of suggestion must generate different language for the suggestion. This is handled using advanced insight.

3. **Suggestion**

 As the suggestion is triggered by on the trigger conditions on HCP Data, corresponding action is determined by the action-engagement model which then decides the action to be taken and that is configured in Action Text.

4. **Dismissal Survey**

These suggestions are pushed for a specific brand and every brand is related to some feature which might not be the same for rest of the brands. This distinguishing is done using a survey id (configured in Dismissal Survey Tab).

5. **Disable Suggestions**

During the run, certain suggestion may prove to be either irrelevant for not useful. This could be settled using disabling this suggestion which won't let it trigger suggestion when the engine runs.

7. SUMMARY

Development for the different products of the client I was working for went smoothly and are performing well. These products are now live and provide suggestions to the reps nationwide in the US every Monday. It's been a successful launch and a good teamwork lesson for me. All the issues faced during migration were duly noted to avoid them next time.

Performance analysis of the brand clearly specifies a jump in profit margin due to effective interaction between HCP-REP. The suggestions provided were reported to be useful and the feedback was duly noted to improve the next cycle of suggestions. The project is being implemented successfully, and the requirements are being satisfied complying with the organizational policies, without major changes to the scheduled timeline. The development process in general was quite smooth and was a good experience for both the organisation as well for me as a developer. The feedback from the organisation and its key players was excellent.

8. REFERENCES

1. https://www.zs.com/services/technology/technology-solutions/ccmtechnology

2. https://www.zs.com/services/technology/technology-solutions/crm-solutions

3. https://www.zs.com/services/software/orchestration-engine

4. https://www.zs.com/services/technology

5. https://aws.amazon.com/ec2/

6. https://aws.amazon.com/s3/

7. https://aws.amazon.com/health/biotech-pharma/commercialization/

Project Report Undertaking

I Mr. /Ms. ___Prashasti Dwivedi _____
-Roll No._____171462_____ Branch_____Information Technology_____ is doing my internship with __ZS Associates_____ from____1 Feb, 2021_____ to __31 July, 2021_____

As per procedure I have to submit my project report to the university related to my work that I have done during this internship.

I have compiled my project report. But due to COVID-19 situation my project mentor in the company is not able to sign my project report.

So I hereby declare that the project report is fully designed/developed by me and no part of the work is borrowed or purchased from any agency. And I'll produce a certificate/document of my internship completion with the company to TnP Cell whenever COVID-19 situation gets normal.

Signature: *Prashasti Dwivedi*

Name : Prashasti Dwivedi

Roll No: 171462

Date : 25-05-2021