



# **Automated Insights for Digital Evolution**

Digital Optimization platform

## Outline

**01** What is A.I.D.E.

**02** Use Case

**03** Data & Infra Requirements

# What is A.I.D.E.?

# How does AIDE improve digital experiences

**AIDE** mines through millions of **digital touchpoints**, along with **external and omni-channel** data to uncover several microscopic factors that causes dissonance using a series of **unique pattern recognition AI algorithms**

## 1. SENSORIZE



Feature engineering across structured, unstructured and semi-structured data domains

## 2.DETECT



Actively monitor digital indicators across the journey to detect issues and opportunities

## 3. LOCATE



Pinpoint granular issues on the website and quickly recover leads and revenue fast

## 4.REVEAL



Voice of Customer analysis to uncover the contexts behind the frictions or struggles

## 5. PRIORITIZE



Measure and prioritize insights based on opportunities or lead potential

Insight as a Service

Platform as a Service

Managed Services

On-Premise | Popular Cloud Platforms

# What is friction and how do we define it

**Missing artifacts, Improper design, Redundant steps, or Process errors** constitutes friction

## MISSING ARTIFACTS

Difficulty in understanding the reasons for an **increase in the bill amount** on digital channel led to a subsequent call seeking explanation for a leading US Telecom provider

**Non-US based address error** while trying to add an authorized user to the Credit Card account led to an increased dependency on Care channel

## ERRORS



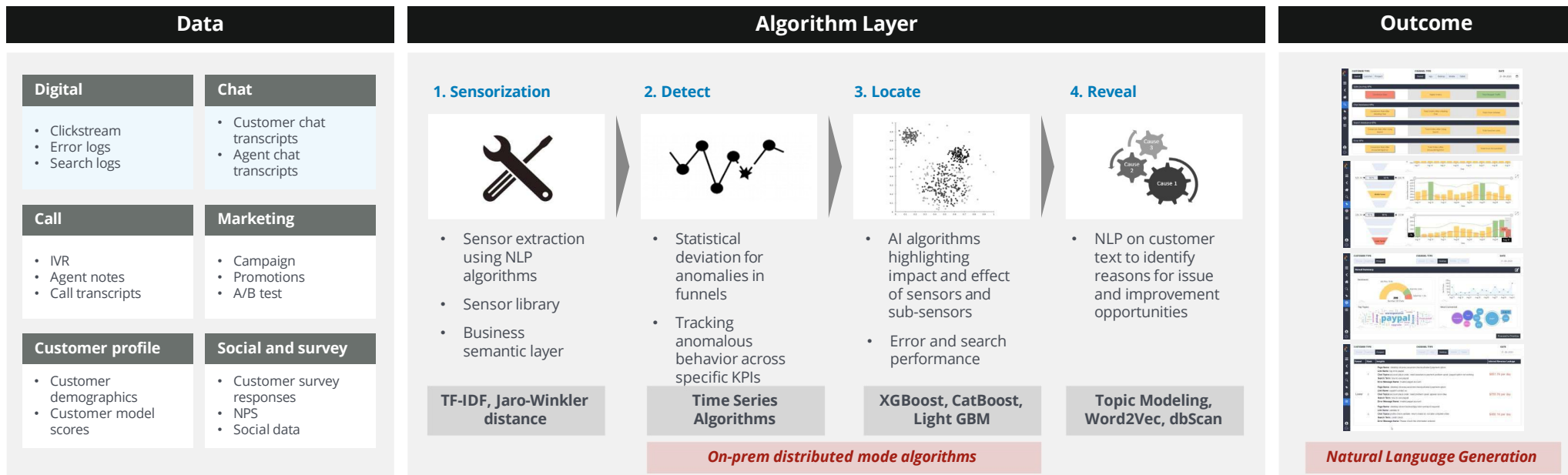
## NAVIGATION

Mortgage business faced **too many steps** for prospects arriving through organic search looking to reprice existing loan, resulted in ultimate drop-off

Website **did not allow a side-by-side comparison of the credit cards offered** which led to ultimate abandonment of the sales journey after printing the page

## DESIGN

# A.I.D.E. uses customizable, open-source AI that work well with the complex journeys, data security and time to market



On-Premise | Popular Cloud Platforms (AWS; Azure; GCP)

# Addressing the “Why” and not just the “What” is how AIDE differentiates from conventional platforms like Adobe

Moving away from **monitoring surface-level metrics** such as clicks, conversions and bounce rates, to a **reason behind** those results is the new way of analyzing clickstream data feed

01

## **True friction v/s Focus on exit point**

“Exits on the page” DOES NOT necessarily mean the visitor faces the friction on that page. Millions of journey touchpoints mined using AI recognize interaction patterns which leads to abandoning the journey

02

## **Cross-channel interaction feed**

Customer interactions before landing on the digital channel and post exiting the digital channel provides a granular level of insights to understand customer’s decision-making behavior

03

## **Context identification**

Uncovering root cause is not possible using only clickstream data feed. Integration with Call, Chat, Store and other unstructured data domains helps to piece together root cause of friction/dissonance

04

## **Data as a Product**

Automated pipeline to derive meaningful business features at scale using the semi-structured clickstream feed to support multiple downstream AI/ML use cases

## **What AIDE is not?**

AIDE is not a tag management platform or a clickstream data provider

# Sample Use Case

# Data & Infra requirements

# How have we enabled data & analytics team across different industries to use clickstream feed at scale

Low

High

Data maturity

## No tag management

A leading **US based home mortgage originator** wanted to **set-up their digital feature store** to drive **personalization initiatives**

Leveraged **GlassBox** to enable **tag-less data** capture for **website(.com)** and **mobile app**

## No data feed available

A leading **US based investment management firm** wanted to **set-up an automated real-time pipeline** to detect anomalies

Leveraged **KSQL Stream Processing** and real-time databases in **Kafka** to set-up and analyze streaming data

## JSON repository

A leading **Telecom provider** wanted to create a **structured data table** on the repository of json files to create a **single source of truth**

**On-prem** data stack was leveraged by Fractal's **BDE team** to create a structured clickstream data table

## Scaled features

A leading **Media provider** wanted to **automate intelligent feature** creation to enable **downstream AI/ML & reporting use cases**

Leveraged **Sensorization** to enable **automated feature scaling** using a combination of **NLP and Engineering**

# Pre-requisites across Data, Software and Infrastructure to enable AIDE implementation

## Data Requirement

- **Raw data**
  - Web Click stream data with high coverage tagging
  - E-Chat raw transcripts (Customer and Agent text)
  - Meta data (OS, Browser, Chat duration etc.)
  - Customer Demographic / Ownership data
- **Data pre-processing**
  - Semantic: Definition of journey stage (U/M/L)
  - Chat & Search event should be mapped to a Page
  - Visitor mapping to end digital product application

## Good to Have Data domains

- |                    |                 |
|--------------------|-----------------|
| ▪ Mobile App feed  | ▪ Survey        |
| ▪ Call transcripts | ▪ Branch Visits |
| ▪ Feedback         | ▪ Campaign      |

## Software Requirement

- **Programming Language**
  - Python: Version 2.7 and above
  - R: Version 3 and above
- **ML / Visualization Libraries**
  - Python: NumPy, SciPy, Scikit-learn, pETL
  - R: MLR, parsnip, purr, Ranger
- **IDE**
  - Jupyter, PyCharm or Rstudio
- **Data Processing**
  - SQL
- **Visualization**
  - Tableau: Version 9.2 and above
  - QlikView: Version 11 and above

## Good to Have Software

- |   |   |
|---|---|
| <b>Programming Language</b> <ul style="list-style-type: none"> <li>▪ Spark: Version 2.4 and above<br/>[Ability to support scale]</li> </ul> | <b>Programming Language</b> <ul style="list-style-type: none"> <li>▪ Kibana: Version 6 and above<br/>[Real time Visualization]</li> </ul> |
|---|---|

## Infrastructure Requirement

- **Assumption**
  - Size of the data in TB with a retention policy of 1 year
    - Container storage: 30% of the data
    - Compressed Parquet format: 70% of the data
- **Storage requirement**
  - Relational databases(anyone) like Oracle, MySQL, SQL Server, Teradata
- **Computation requirement**
  - 100+GB RAM and 32 CPU core

## Good to Have Infrastructure

- |  |  |
|--|--|
| <b>Number of Data nodes</b> 325 with each disk worth 20 TB and OS memory of 4 GB | <b>Batch processing:</b> 3000 core processors<br><b>In-memory processing:</b> 4000 core processors |
|--|--|

# Thank You.



A strategic partner to the most admired Fortune 500® companies globally, we help power every human decision in the enterprise by bringing advanced analytics & AI, engineering and design.



[www.fractal.ai](http://www.fractal.ai)