**NRG Business Sales & Marketing Data Enablement**

MEETING DETAILS

|  |  |
| --- | --- |
| **DATE** | **MEETING TITLE OR PURPOSE** |
| **01/15/25** | **Data & AI Team Initial Discovery** |
| **ATTENDEES** | |
| * Richard Wang – Director, Data & AI * NRG Team   + Felix Chik   + Andrew Eckenrode   + Jessica Paz * Credera Team   + Taylor Gregston   + Luke Hamilton | |

MEETING SUMMARY

|  |
| --- |
| **NEXT STEPS** |
| * **01/15:** Luke to share slides and remaining questions with Richard for review * **01/15:** Richard to introduce project to Vincent tomorrow and get initial perspective * **01/15:** Richard to share feedback from Vincent and indicate timing/attendees for next meeting |
| **KEY TAKEAWAYS** |
| * Marketing/CX data for prospects and existing customers is expected to be included in the foundational model * Open questions remain around ownership of development effort and location for buildout of model. We will need to work with Vincent to determine the preferred approach * Vincent's team has many projects planned for early 2025; we will need to get this prioritized on the roadmap to keep our current timeline |

FULL MEETING NOTES

|  |
| --- |
| **RAW NOTES** |
| Data Architecture Preferences:   * Overview of existing data architecture in place (e.g., Data Lake, IRW tables, data warehouses)   + What is the current state of the Data Lake?     - Has connection to IRW tables. Already have a lot of data sources, looking to continue growing the data lake     - In progress Clara (customer care chat bot) project – adding Siebel, SAP data further to address customer care actions     - Desired approach – avoid connection with source systems directly; pull data from data layer systems where possible     - UCV – unified customer view – is the path that the data team is taking moving forward       * Logical model is being created to transform data. Will also have associated unstructured data       * Vincent’s team may be able to provide more detail on available data     - Current documentation of available data in data lake – there is so much data in the data lake that it is difficult to share everything that is available. Rather, it is easier to ask about specific data for the data team to check if it is available     - MDM – master data management – project in progress. Targeting late 2025 to have this complete. Purpose is to create a common ID to find all correlations between systems     - Data team development – generally, the data engineering team prefers to own development of data lake initiatives       * Separate data warehouse development efforts still need to conform to data governance guidelines       * Need to discuss further with Vincent’s team to understand development, architecture, and location following discovery efforts   + What data warehouses already exist that can be leveraged?   + IRW tables – is this data already available in a data lake/warehouse, or only on-prem? * Confirm assumption that AWS is still the preferred cloud provider   + Data lake is moving from AWS to GCP in late January-early February   + Preferred provider will be GCP moving forward * Data warehousing – what products are used for data warehousing in the current environment?   + Is AWS Redshift used for data warehousing? * ETL – what products are used for ETL in the current environment?   + Is AWS EMR used for ETL? * Location for creation of Foundational Model for this project * What environments are used for existing data systems (e.g., dev, stage, and prod)   Project Management and Version Control:   * Where should development tickets related to this effort be created? (Business Solutions, Digital IT Projects, or another ADO project?) * What repositories are needed for the preferred architecture (ADO repo vs other)? * What CI/CD pipelines are needed for the preferred architecture (ADO, Jenkins, etc.)? * Who should we work with for provisioning of required data warehouse, ETL pipelines, repos, and CI/CD pipelines? * Discuss ownership of technical ownership of foundational model following Credera roll off   Documentation and Reporting:   * Confirm that ADO Wiki is still the preferred location to store technical documentation * Is there a preferred diagramming suite used by the Data & AI teams that we can request access to?   + Ideally, this would integrate with the ADO wiki or other documentation suite. If not, static images will be utilized and files can be shared once complete * Power BI – output from Foundational Model will utilize Power BI. Are there any additional permissions needed to create new Power BI reports?   Source Systems:   * Data lake – what data is available in the data lake that could be leveraged for this project? * On-Premises Oracle DB – understand what data is still stored on-premises vs being migrated to the cloud/data lake |