Requirements Document

Uday Jain

AI Traffic Forecast

PA2578 HT20

VERSION: 1.5 REVISION DATE: 10th May 2024

Contents

[Section 1 Introduction 3](#_gjdgxs)

[1.1](#_30j0zll) Purpose and scope 3

[1.2](#_1fob9te) Definitions, acronyms, and abbreviations 3

[1.3](#_3znysh7) Overview 4

[1.4](#_2et92p0) Goals of the product (goal level requirements) 4

[1.5 Context diagram for the system](#_tyjcwt) 4

[Section 2 Stakeholder Identification and analysis](#_3dy6vkm) 5

[Section 3 Requirements Elicitation Techniques](#_1t3h5sf) 7

[3.1](#_4d34og8) Elicitation Technique 1 (Interviews) 7

[3.2](#_2s8eyo1) Elicitation Technique 2 (Brainstorming) 7

[3.3](#_17dp8vu) Elicitation Technique 3 (Requirements Workshop) 8

[Section 4 System Requirements](#_3rdcrjn) 8

[2.1](#_26in1rg) Domain Level Requirements 8

[2.2](#_lnxbz9) Functional Product Level Requirements 9

[2.3](#_35nkun2) Data Requirements 9

[2.4](#_1ksv4uv) Product Quality Requirements 10

[Section 5 Requirements prioritization](#_2jxsxqh) 11

[Section 6 Policy and Regulation Requirements](#_1y810tw) 12

[Section 7 References](#_4i7ojhp) 13

[Section 8 Document Revision History](#_2xcytpi) 13

[Section 9 Appendices](#_1ci93xb) 13

# Section 1 Introduction

* 1. Purpose and scope

The product is envisioned to help common & country product managers, service managers, solution managers, business managers for mobile services, network & RAN engineers as well as country specific operational teams such as sales and marketing take informed decisions to optimize operations and improve user experience through network planning, marketing and sales using accurate and reliable forecasts for data usage and user volumes.

The AI Traffic forecast product would provide accurate total data usage and subscriber count monthly forecasts for 36 months in future across multiple hierarchical levels of municipality, product type, subscription type, technology type and spectrum band.

The current scope of product excludes un-classified product categories or un-identified users (contributing to <X% of total data usage during 2023).

The forecasts would be accompanied by confidence intervals that would help the users take informed decisions in an uncertain future.

The forecasts do not account for sudden changes in sales strategy or major deviations in market sentiments in the short term. However, the forecasts are expected to learn the changes in trend during regular refresh cycles.

* 1. Definitions, acronyms, and abbreviations

Provide the definitions of all terms, acronyms and abbreviations that you have used in the document. As a suggestion, you could use a table. Do not forget to specify here how you refer to each requirement (identity).

If you have any domain specific terms or domain descriptions please add them here.

| **Acronyms & Abbreviations** | **Definition** |
| --- | --- |
| AI | Artificial intelligence |
| RAN | Radio access network |
| MAPE | Mean absolute percentage error |
| System 1 (name hidden owing to data privacy control) | Internal system capturing historical data usage data |
| System 2 (name hidden owing to data privacy control) | Internal system capturing historical subscriber details |
| System 3 (name hidden owing to data privacy control) | Internal system capturing historical competitor details |
| System 4 (name hidden owing to data privacy control) | Internal system capturing historical pricing and revenue details |
| System 5 (name hidden owing to data privacy control) | Internal system capturing historical country level macro-economic data |
| System 6 (name hidden owing to data privacy control) | Internal system capturing historical infrastructure data |
|  |  |

* 1. Overview

The document provides a detailed overview of requirement gathering and prioritization exercise done for AI traffic Forecast product starting with stakeholder and data source identification.

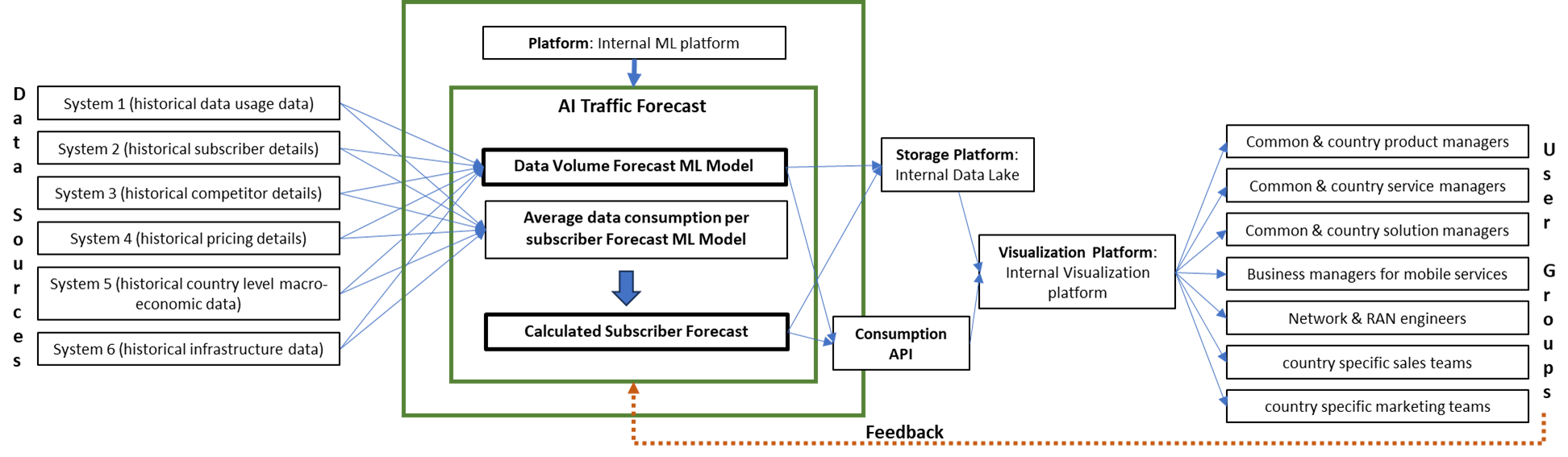
## Goals of the product (goal level requirements)

Goal\_1: Forecast capacity across different geographical zones

Goal\_2: Highlight areas with mis-managed capacity utilization for sales and marketing functions as well as network operational teams

Goal\_3: Forecast models should have a MAPE <15%

## 1.5 Context diagram for the system (please insert the context diagram here)



# Section 2 Stakeholder Identification and analysis (see lecture 2 about Elicitation)

This section lists the stakeholders for the system you are developing. List the stakeholders here, describe them briefly and categorize them according to the groups of interest or importance.

| **Stakeholder** | **Description** | **Category** | **Score (0-5)**  (higher is better) |
| --- | --- | --- | --- |
| Common & country product managers | Product managers own the roadmap for different product portfolios. These are key potential stakeholders that may use forecasts to analyze growth and take strategic product decisions | Internal stakeholder who can also provide value from experience as well as strategic direction. Provide influence. | Expertise: 5  Influence: 5  Behavioral: 2 |
| Common & country service managers | Service managers own the service metrics and ensure consistent user experience. These are potential stakeholders who would use forecasts to analyze performance of respective domains and plan to minimize any significant deviations in customer experience due to overcrowding and over capacity utilization | Internal stakeholder who can also provide value from customer/market understanding as well as strategic direction. Provide influence.  Also potential collaborators. | Expertise: 4  Influence: 5  Behavioral: 5 |
| Common & country solution managers | Solution managers own multiple epic portfolios for different product areas. These are potential stakeholders who would use forecasts to analyze performance of respective initiatives and plan for strategic growth. | Internal stakeholder who can also provide value from domain expertise as well as strategic inputs. Provide influence. | Expertise: 4  Influence: 2  Behavioral: 3 |
| Common & country business managers for mobile services | Business managers own a portfolio of mobile services for different products and geographical areas. These are potential stakeholders who would use forecasts to analyze performance of respective portfolio and plan for strategic growth. | Internal stakeholder who can also provide value from domain expertise as well as business processes. Provide influence. | Expertise: 3  Influence: 5  Behavioral: 4 |
| Network & RAN engineers | Network & RAN engineers continuously monitor low level network performance metrics and ensure consistent user experience. These are potential stakeholders who would use forecasts to analyze performance and plan to minimize any significant deviations in customer experience due to overcrowding and over capacity utilization | Internal stakeholder who can also provide value from technical expertise and common practices that influence observed data | Expertise: 4  Influence: 0  Behavioral: 0 |
| Country specific sales team | Sales teams own a portfolio of products and services for different geographical areas. These are potential stakeholders who would use forecasts to analyze performance of respective portfolio and plan for strategic growth within their respective domains to improve growth. | Internal stakeholder who can also provide value from customer/market understanding as well as strategic direction through planned strategies. Provide influence.  Also potential collaborators. | Expertise: 3  Influence: 2  Behavioral: 5 |
| Country specific marketing team | Marketing teams own a portfolio of marketing initiatives (brand and action campaigns) for different geographical areas. These are potential stakeholders who would use forecasts to analyze unused capacity and plan for strategic growth. | Internal stakeholder who can also provide value from customer/market understanding as well as strategic direction through planned strategies. Provide influence. Also potential collaborators. | Expertise: 1  Influence: 1  Behavioral: 4 |

| **Data Source** | **Description** | **Category** | **Score (0-5)**  (higher is better) |
| --- | --- | --- | --- |
| System 1 | Internal system capturing historical data usage data | Low cost, easy to access and high data availability system that provide accurate data | Cost: 5  Easy of access: 5  Availability: 5  Accuracy: 5  Opportunity cost: 5 |
| System 2 | Internal system capturing historical subscriber details | Low cost, easy to access and medium data availability system that provide accurate data | Cost: 5  Easy of access: 4  Availability: 4  Accuracy: 5  Opportunity cost: 5 |
| System 3 | Internal system capturing historical competitor details | low data availability system that comes with high opportunity cost | Cost: 2  Easy of access: 2  Availability: 1  Accuracy: 2  Opportunity cost: 5 |
| System 4 | Internal system capturing historical pricing and revenue details | Low cost, easy to access and high data availability system that provides medium accurate data | Cost: 5  Easy of access: 1  Availability: 5  Accuracy: 3  Opportunity cost: 2 |
| System 5 | Internal system capturing historical country level macro-economic data | high cost and low data availability system that comes with high opportunity cost | Cost: 2  Easy of access: 3  Availability: 2  Accuracy: 1  Opportunity cost: 5 |
| System 6 | Internal system capturing historical infrastructure data | Low cost, easy to access and high data availability system that provide low accuracy data | Cost: 5  Easy of access: 5  Availability: 5  Accuracy: 3  Opportunity cost: 3 |

# Section 3 Requirements Elicitation Techniques

Since the product is a ML model, no mock-ups or visual prototypes were discussed during any of the requirement elicitation sessions.

## 3.1 Elicitation Technique 1: Interviews:

## To kick start the process interviews were done with representatives across the stakeholder groups. These interviews were kept unstructured to capture issues and current processes in detail, helping us formulate requirements. Each interview was done 1x1 within a 30 min slot either virtually or in-person when possible. The following set of requirements were formulated based on these sessions:

1. Goal\_1: Subscriber and data forecast are required not only at country level but also for each geographical zone/municipality.
2. Goal\_2: Data forecast combined with installed capacity can help teams identify areas with additional capacity as well as areas with need of additional capacity.
3. Current planning is done ad-hoc using simple linear regression based models coupled with intuitions that often suffer from lower accuracy and are difficult to reconcile.

## 3.2 Elicitation Technique 2: Brainstorming:

## Post initial discussions and literature review we did a whiteboard brainstorming session to identify potential solutions and factors associated with forecasting subscriber and data growth. The session was conducted with a key set of stakeholder representations selected from the pool to form a focus group.

## The following list captures key outcomes:

1. Forecast split across customer type, subscription type as well as technology type would add value and help multiple other use-cases
2. Goal\_3: To help teams with operational team plan sales, marketing campaigns or capacity a MAPE<=15% would be sufficient.
3. Identified System 1-6 as key sources that can help capture factors impacting subscriber and data growth.
4. Owing to data privacy and quality issues with multiple data sources, it may be difficult to source all the data in the first stage itself.

A separate brainstorming session with Data Science practice helped identify different modeling methodologies and workflows that can be explored to help design the solution.

## 3.3 Elicitation Technique 3 : Requirements Workshop:

## Post multiple discussions and inputs from the community we consolidated the list of requirements to plan for the product roadmap. The findings and initial draft of the roadmap was discussed with the focus group to gain consensus as well as chalk out the intricate details for the plan and resolve prioritization conflicts.

## This exercise was instrumental in helping us identify top priorities as well as plan deliverables/release cycles.

The following lists some of the key decisions made during the workshop.

1. A monthly forecast spanning across 3 years in future would be required to provide sufficient time to teams for planning and relevant actions.
2. First product cycle would include factors only from data sources 1,2 and 6.
3. Second release would include factors using data source 3.
4. Future releases post version 2 would depend on feedback, adoption and accuracy levels

# Section 4 System Requirements

To capture requirements following techniques were used:

1. Data model : Illustrated below in section 4.3. Owing to data privacy issues the actual data columns have been hidden.
2. Data dictionary : Owing to data privacy issues it was difficult to provide any specific details
3. Feature requirements See individual Functional Product Level Requirements.
4. Vivid Scenario: This technique was used to capture different scenarios where the model would exclude/include in prediction. eg.

Maintenance of 4G Tower for in Municipality X: During the week 24, tower XXX in municipality X was shut down and traffic was rerouted to nearby 3G towers. This led to a major jump in data consumed on 3G in municipality X and a consequent drop in 4G data consumption. Since this is a short term event, the model does not learn the behavior and excludes sudden changes in 3G/4G data consumption for municipality X.

Summer season: Every year Swedish residents move from metropolitan cities such as Stockholm to smaller towns and remote geographical areas during the months July and August. Since this trend is observed across all years, the model learns the seasonal behavior and forecasts lower data consumption during July and August for big cities while higher data consumption for rural and otherwise sparsely populated geographical zones.

1. Context diagram: See section 1.5

## 4.1 Domain Level Requirements

DL1: Model should account for seasonal changes in data consumption trends across municipalities and other geographical zones. eg. major subscribers move to rural/semi-urban municipalities during the summer season.

DL2:Model should adapt to changes in sales strategy and learn the growth cycle of different products over time within 3 months of major changes. While 3G is being phased out, 4G is saturated and 5G is showing growing adoption.

DL3: Model should account for capacity changes and maintenance plans across different geographical zones. Short term changes spanning less than 1 month can be excluded, however long term changes to capacity should be identified and included as part of the prediction.

## 4.2 Functional Product Level Requirements

PR1: Forecast ML model runs each month to forecast values for subscriber count and data consumption for the 30th Month in future.

PR2: Model refresh or update occurs every X months to ensure unforeseen changes in market conditions are accounted for.

PR3: Model should forecast subscriber count and data consumption for each combination of dimensions.

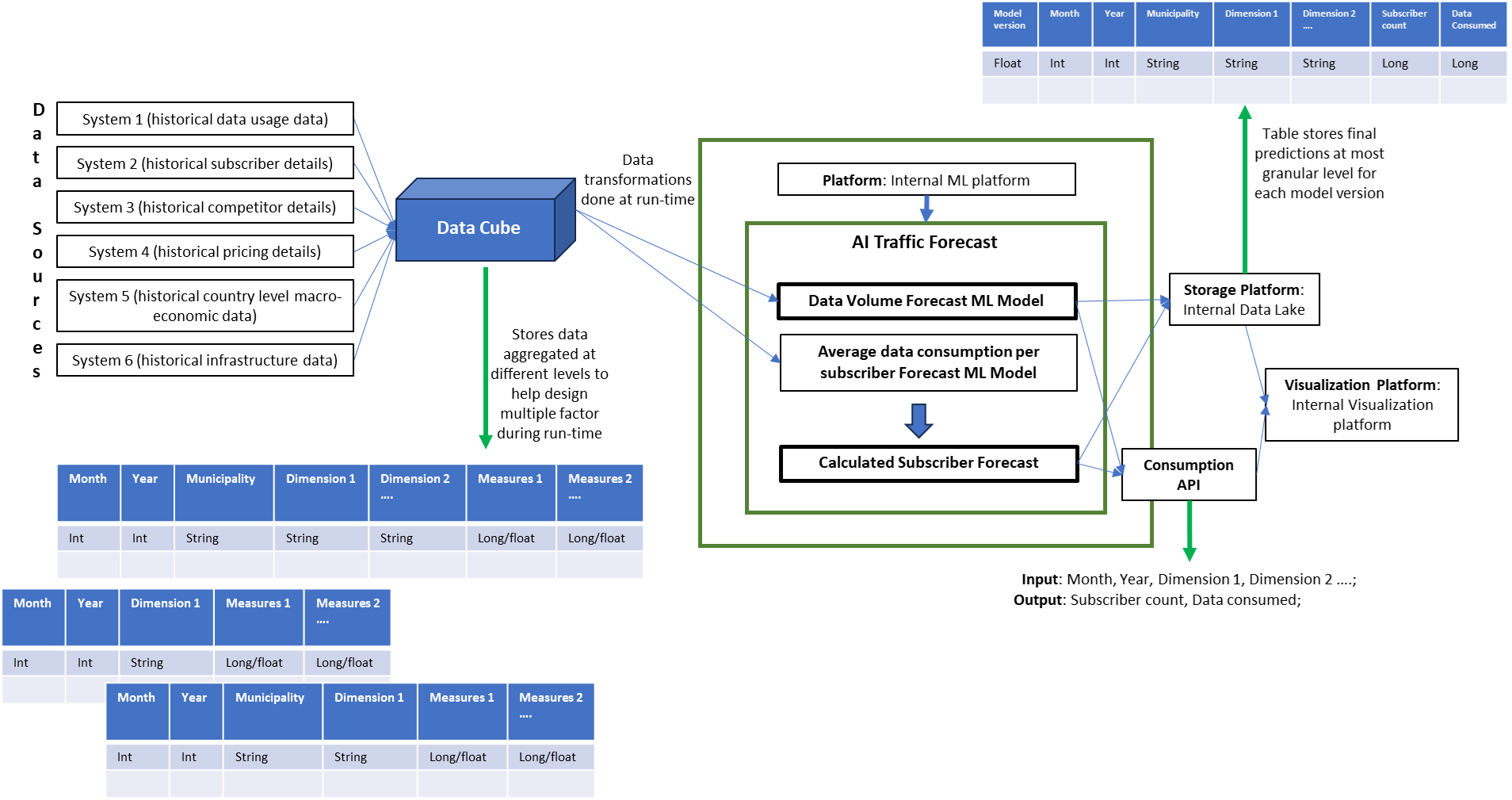
PR4: Model should forecast subscriber count and data consumption for each municipality.

PR5: Forecasts should be accompanied by respective confidence intervals to ensure ease of use and reliability.

## 4.3 Data Requirements

Owing to data privacy issues the actual data columns have been hidden.

The tables represent views/tables with relational data model.



DR1: Data cube contains measures aggregated at multiple levels of dimensions. The levels should cover all possible permutations of different dimensions for each month eg. all measures aggregated for each municipality for each month, all measures aggregated for each municipality and categories of dimension 1 for each month.

DR2: Values of month, year along with combination of dimensions in specified order needs to be passed to API to get predicted subscriber count and data consumed predictions in return.

DR3: Prediction data is stored in a table using a rational data model with prediction values for each month and combination of dimensions to most granular level eg. predicted subscriber count for May 2025, Stockholm municipality, dimension 1 = X, dimension 2= X2 …

## 4.4 Product Quality Requirements

## Here you specify your quality requirements (also called non-functional requirements). You are expected to write requirements about a minimum of 5 quality aspects (e.g. performance, usability etc..)

QR1: Correctness: Model performance statistics such as MAPE are monitored every month to ensure SLA of overall MAPE <=15% is met.

QR2: Usability: 80% of users shall find the visualization forecast dashboard easy to use. 60% shall recommend the dashboard to others.

Visualised Forecast should be easy to understand and be displayed in units commonly used. eg. GB/TB of data consumed per month or X hundred thousand subscribers

QR3: Interoperability: Model should be built using kubeFlow pipeline to ensure easy transition to any system for processing. eg. GCP, AWS, Azure cloud or on-prem data center.

QR4: Integrity: Only internal infrastructure should be used in development to ensure high levels of security and data privacy protocols are followed.

QR5:Portability & Maintainability: Model refresh and update should not take more than 1 sprint cycle.

Code should follow kubeflow pipeline structure laid out as part of best practices for data science development. All functions should be modular and parameterized to ensure easy portability.

QR6: Usability: Forecast visualization should load within 20 seconds and handle a load of at least 10 concurrent users at peak usage.

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# Section 5 Requirements prioritization

We started with the 100-dollar test to estimate the top of mind and contrasting views among stakeholders. The prioritization exercise was done with representatives within the focus group only. Also to note the stakeholder groups may not be equally distributed. The following table represents average score from each of the stakeholder groups:

|  | Stakeholder Group | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Requirement | Common & country product managers | Common & country service managers | Common & country solution managers | Common & country business managers for mobile services | Network & RAN engineers | Country specific sales team | Country specific marketing team |
| DL1 | 35 | 15 | 20 | 30 | 10 | 25 | 30 |
| DL2 | 20 | 25 | 20 | 25 | 20 | 25 | 25 |
| DL3 | 5 | 15 | 20 | 15 | 30 |  |  |
| PR1 | 10 |  |  |  |  |  |  |
| PR2 |  |  |  |  |  |  |  |
| PR3 |  | 10 | 10 | 15 |  | 20 | 20 |
| PR4 | 10 | 15 | 15 |  | 15 |  |  |
| PR5 | 10 | 10 | 5 | 5 |  | 5 | 5 |
| DR1 |  |  |  |  |  |  |  |
| DR2 |  |  |  |  | 10 | 5 | 5 |
| DR3 |  |  |  |  | 5 |  |  |
| QR1 | 5 | 5 | 10 | 10 | 10 | 20 | 15 |
| QR2 | 5 | 5 | 5 |  |  |  |  |
| QR3 |  |  |  |  |  |  |  |
| QR4 |  |  |  |  |  |  |  |
| QR5 |  |  |  |  |  |  |  |
| QR6 |  |  |  |  |  |  |  |

Since some of the requirements are correlated we found this exercise could have been optimized by grouping some of the dependent requirements. We also found the stakeholder prioritization to be biased towards the business outcome and output as opposed to prioritizing any of the platform/data related requirements that may be crucial to the success of the product.

To simplify the prioritization we clustered the requirements into groups (Numerical Assignment) and did a self ranking exercise based on the delivery version to be presented to stakeholders for approval. Since the ranked release version was influenced by the 100-dollar prioritization exercise, it was well received by the focus group. However, I found the group to lay more focus on v1 and v2 deliveries as opposed to reviewing all versions since later releases might be far in future and other feature requirements may be identified based on the feedback and adoption.

| Requirement | Group | Ranked Release Version |
| --- | --- | --- |
| DL1 | Critical | v1 |
| DL2 | Critical | v1 |
| DL3 | Standard | v2 |
| PR1 | Standard | v1 |
| PR2 | Optional | v2 |
| PR3 | Standard | v2 |
| PR4 | Critical | v1 |
| PR5 | Standard | v1 |
| DR1 | Critical | v1 |
| DR2 | Critical | v1 |
| DR3 | Critical | v1 |
| QR1 | Standard | v2-3 |
| QR2 | Standard | v1 |
| QR3 | Standard | v2 |
| QR4 | Critical | v1 |
| QR5 | Optional | v4 |
| QR6 | Optional | v4 |

# 

# Section 6 Policy and Regulation Requirements

1. Any combinations with less than 5 subscribers are anonymised even after aggregation
2. QR4: Integrity: Only internal infrastructure should be used in development to ensure high levels of security and data privacy protocols are followed.

# Section 7 References

| Document Number | Title | Date | Author/Owner |
| --- | --- | --- | --- |
| 1.1 | Data Dictionary for data source 1 | 16th April 2024 | Data platform team |
| 1.2 | Data Dictionary for data source 2 | 16th April 2024 | Data platform team |
| 1.3 | Data Dictionary for data source 3 | 17th April 2024 | Data platform team |
| 1.4 | Data Dictionary for data source 4 | 17th April 2024 | Data platform team |
| 1.5 | Data Dictionary for data source 5 | 17th April 2024 | Data platform team |
| 1.6 | Data Dictionary for data source 6 | 17th April 2024 | Data platform team |
| 1.7 | ML infrastructure and components | 18th April 2024 | Data Science platform team |
| 1.8 | Network planning and operational maintenance SOPs | 28th April 2024 | Network & RAN Team |

# Section 8 Document Revision History

| Version | Date | Name | Description |
| --- | --- | --- | --- |
| v1.0 | 15th April 2024 | Uday Jain | Added scope |
| v1.1 | 21st April 2024 | Uday Jain | Added stakeholder and data source elicitation |
| v1.2 | 25th April 2024 | Uday Jain | Added details for different elicitation techniques used |
| v1.3 | 5th May 2024 | Uday Jain | Added system requirements and techniques used |
| v1.4 | 9th May 2024 | Uday Jain | added results from requirement prioritization |
| v1.5 | 10th May 2024 | Uday Jain | Added goal |

# Section 9 Appendices