Carbonis
Manage Now For Future

Real-time Carbon Neutrality Management And Optimization Using Natural Language Processing

Project ID: 2022-175

This is our team

Ms. Anjali Gamage	Ms. Sanjeevi Chandrasiri
Supervisor	Co-Supervisor

Sathees P. IT19052748	Mathanika M. IT19005218	Vishakanan S. IT19001562	Vithursan M. IT19033174
Team Leader	Team Member	Team Member	Team Member

Introduction

What is Carbon Accounting?

- **Calculating emission** values for emission activities carried out.
- **Creating various reports** for the different periods.
- **Balance** with available credits.

Our Research Focus & Gap

- How to do carbon reporting in **real-time**?

Research Problem

Questions

1. How can we collect emission activity data efficiently for real-time accounting?
   - Collect from employees using natural language
2. How can we calculate emissions for the emission activities with efficiency?
   - Find and rank emission factors
3. How can we make sure the units are matching in the calculation?
   - Verify and convert before calculating
4. How can we optimize emissions and make sure they achieve those optimizations?
   - Create optimization and send alerts of violations

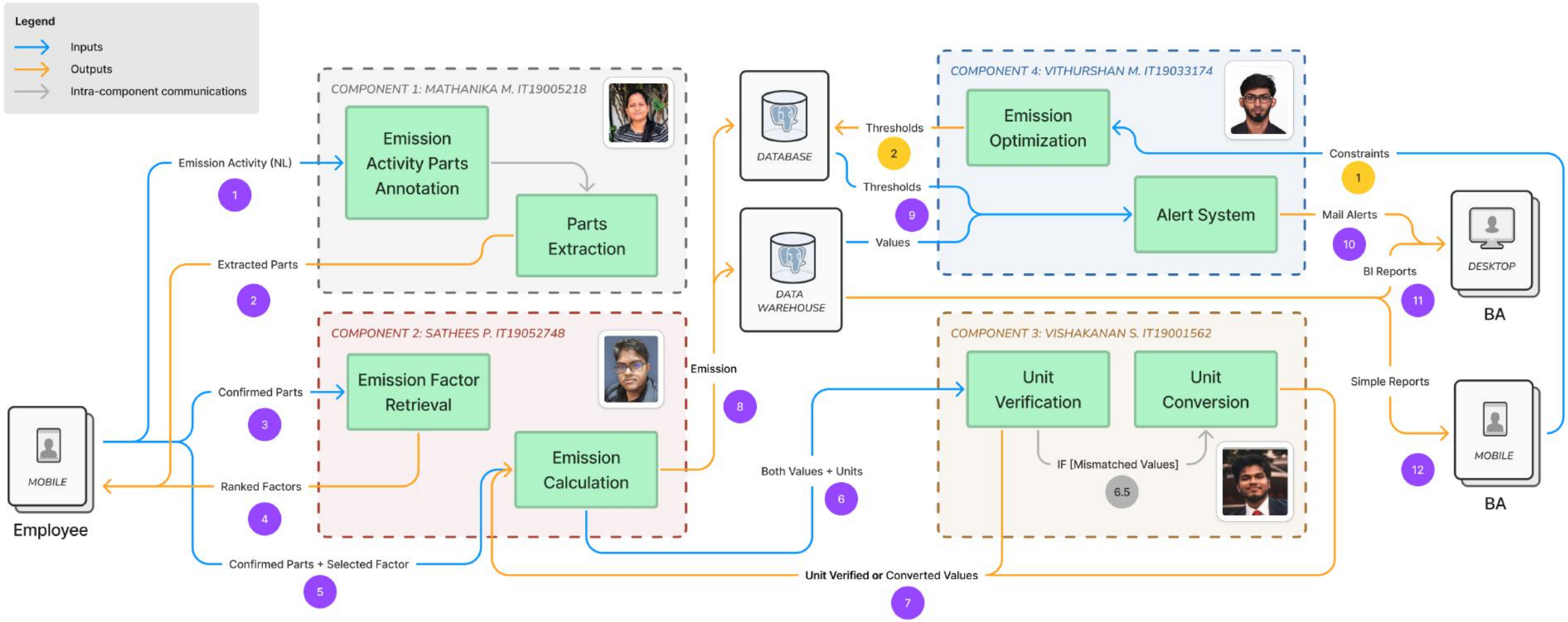
Research Objectives

Main Objective

Create a **cross-platform mobile application** platform for organizations to **manage** and **optimize** their carbon emissions.

Specific Objectives

- Gather employee emission activity details from employees using **natural language**.
- **Search emission factors** and provide **ranked results** for the emission details gathered.
- **Verify and convert values for units** provided by the employees to match the units of the selected emission factor.
- Identify the **optimum solution** for the given emission source constraints and alert about any violations of the optimal solution.



Overall System Architecture (Simplified)

Carbonis
Manage Now For Future

Component 1

• • •

Emission Activity Parts Extraction using Natural Language Processing

Mathanika M.
IT19005218
Data Science

Research Questions

Questions

1. How to gather emission activity data in real-time from employees?

   Natural language input

2. How to identify the emission activity parts?

   Custom named entity recognition

Objectives

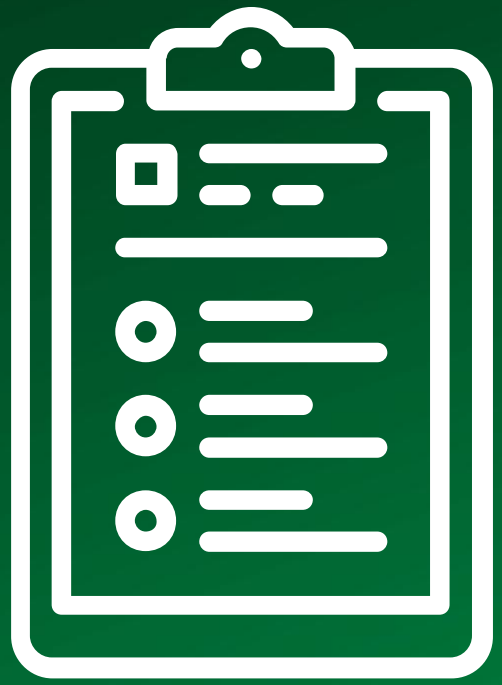
Main Objectives

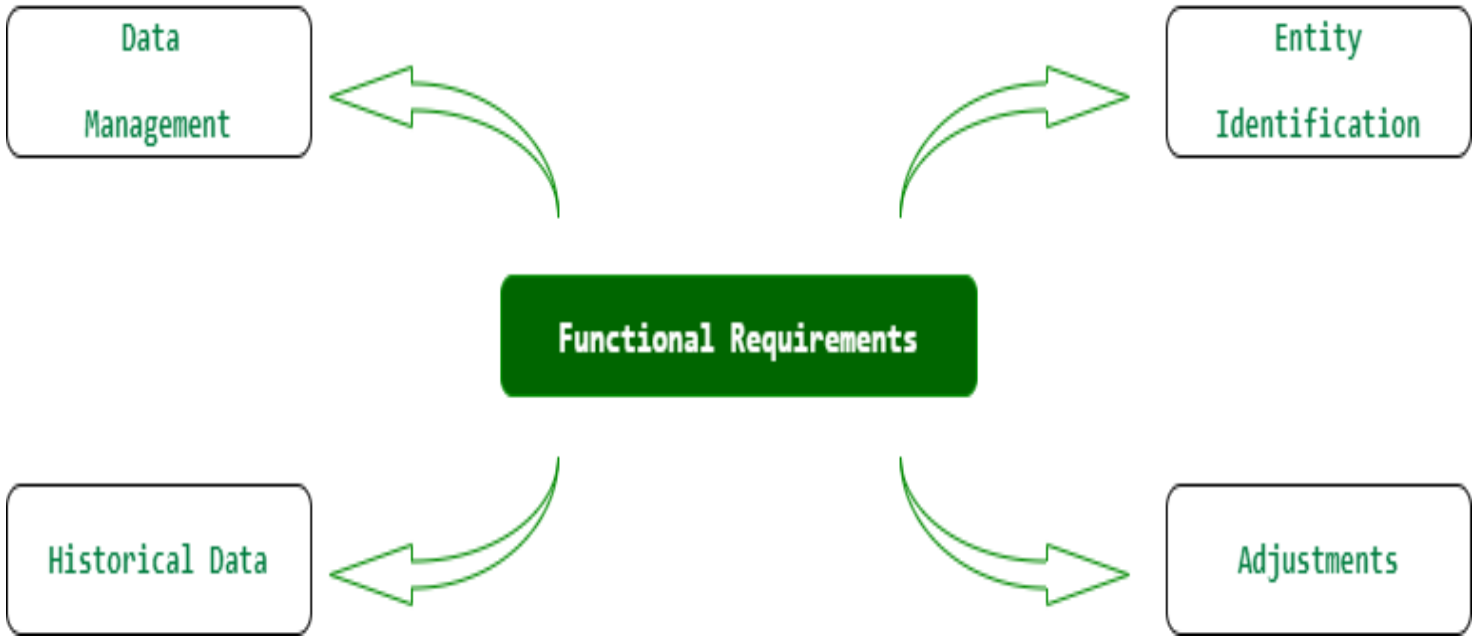
- Collect the real – time emission activity data from the employees using **natural language input**

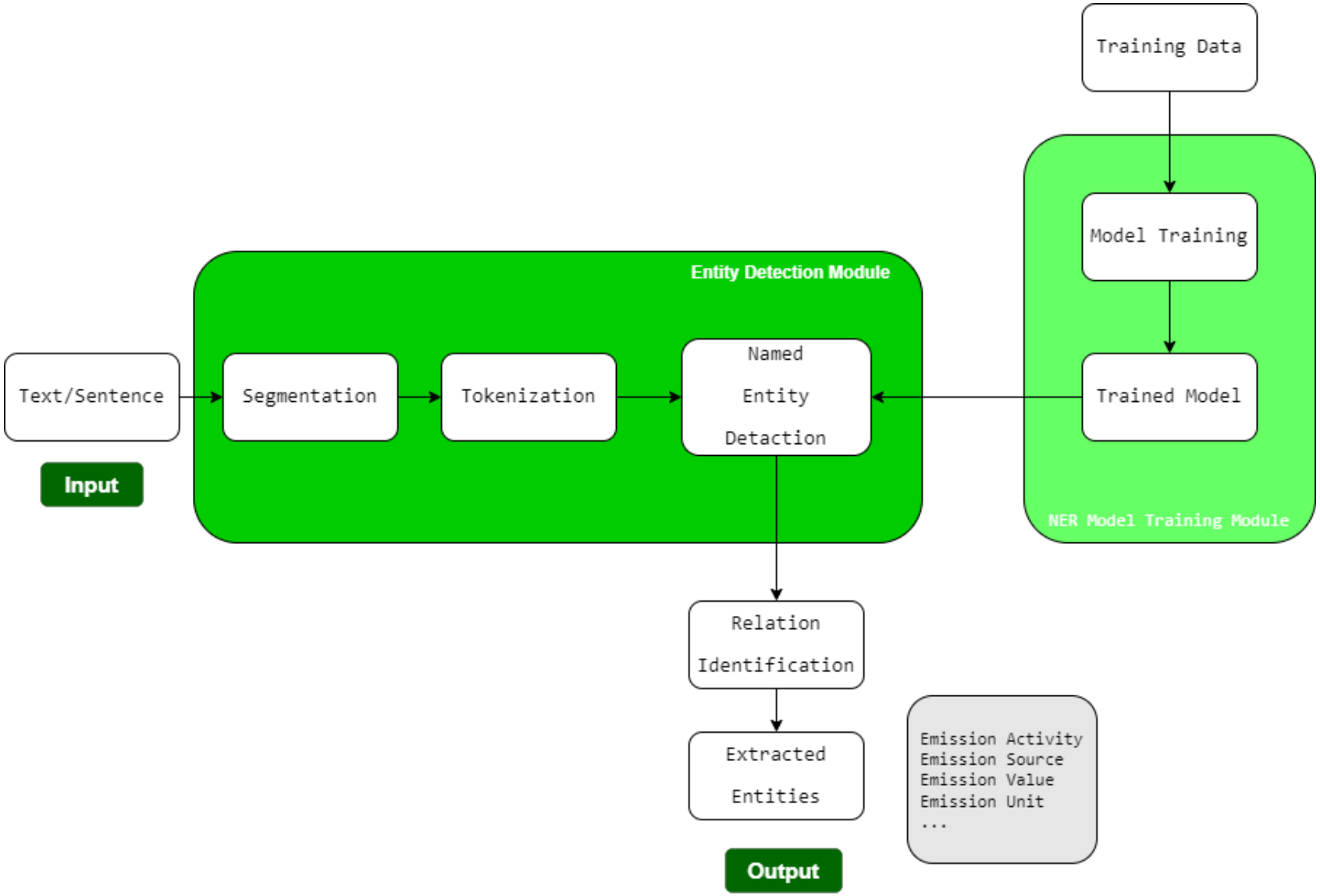
Specific Objectives

- Data collection using natural language
- Data annotation for **custom NER**
- Extraction of emission activity parts

Functional Requirements

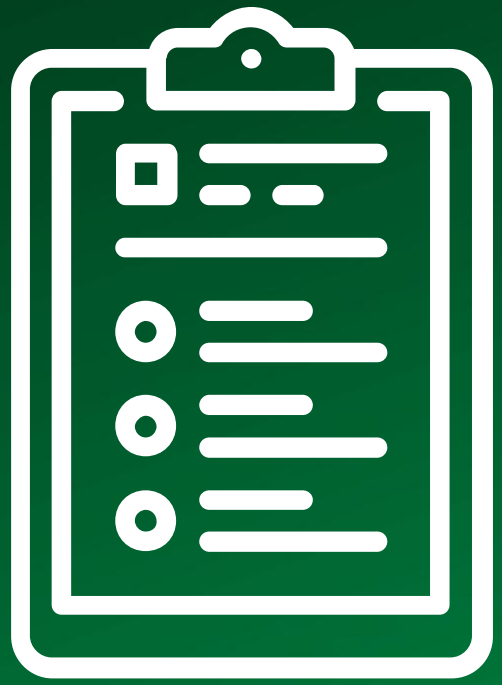


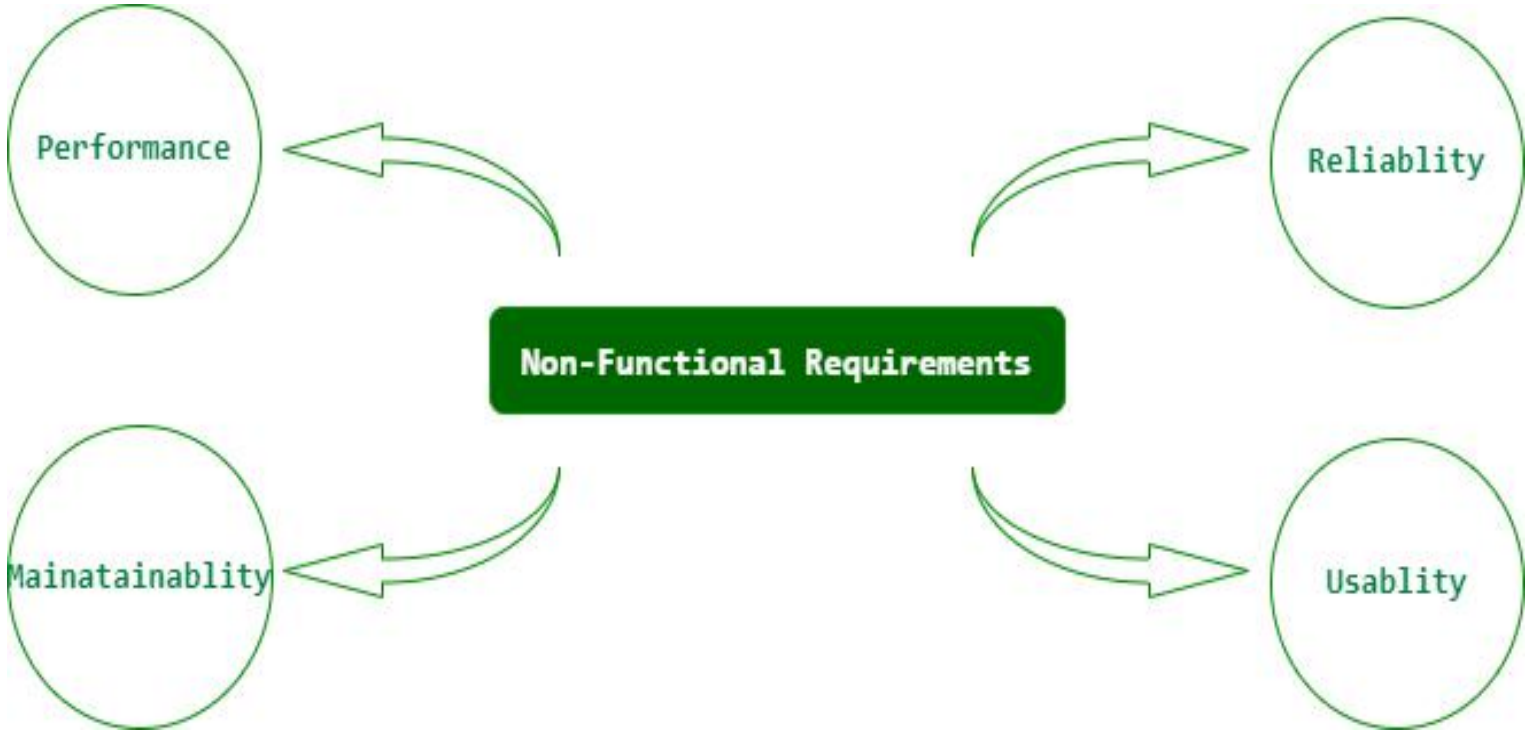




Component 1 Architecture

Non-Functional Requirements

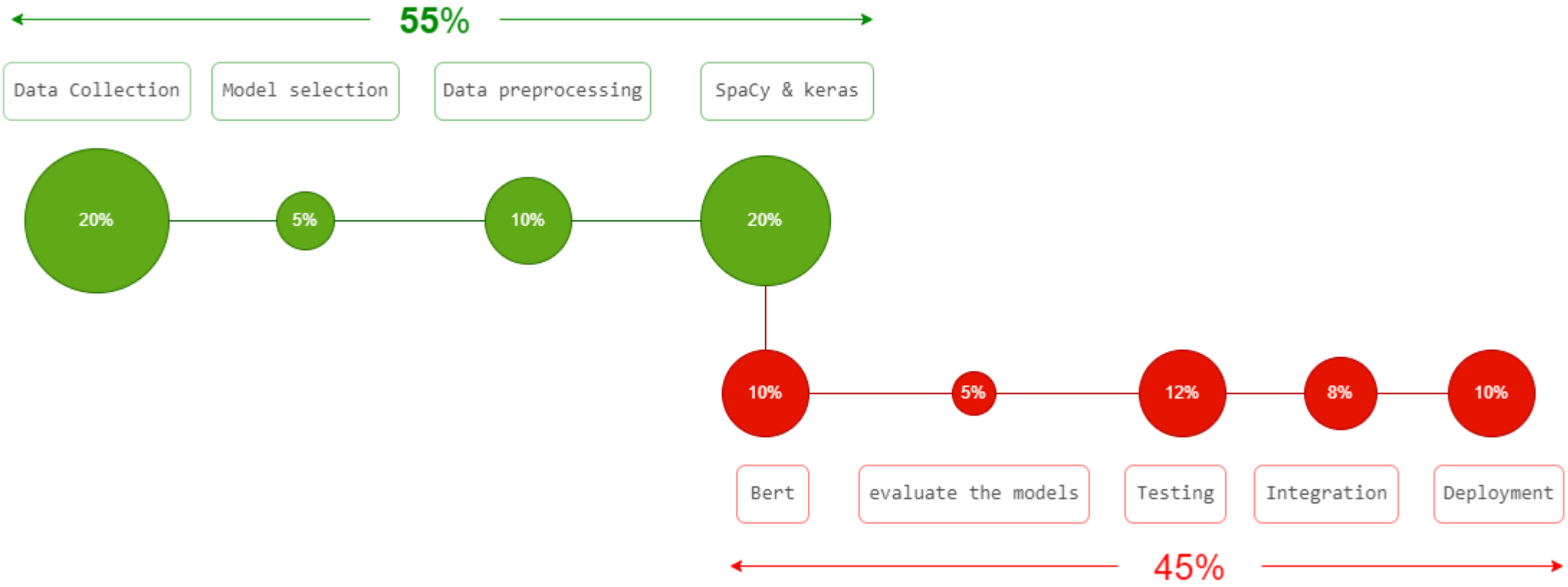




Current Progress – IT19005218

Completed Tasks

1. Data collection
   - data collection through the survey
   - manual data collection
2. Research on model selection
   - SpaCy, keras, BERT
3. Data preprocessing
   - annotation
3. Implementation of models
   - spacy, keras
5. Evaluation of the completed models

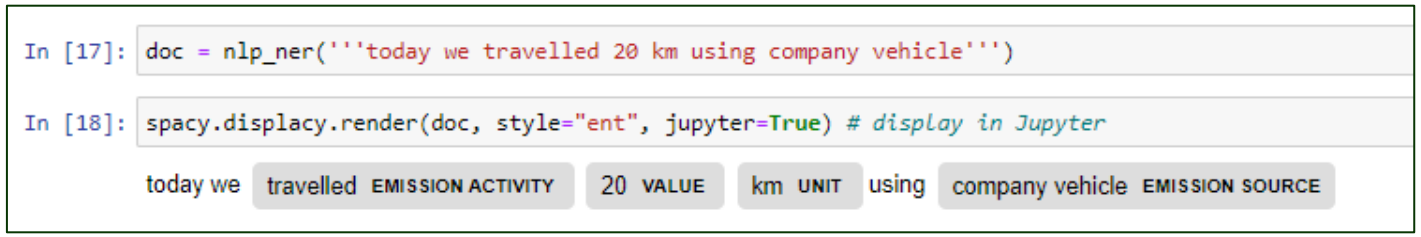


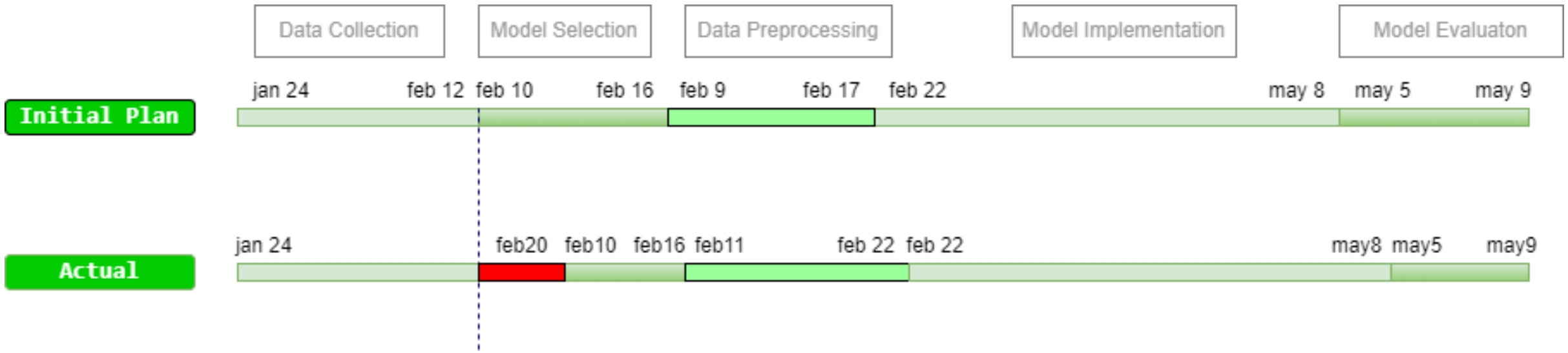
Objectives and Project Completion

Progress Demo (50%)

Carbonis
Manage Now For Future

1. **Proof of concept**
2. **Key pillars of the component 1**
   - Data Annotation
   - Natural Language Processing (NLP)
3. **Technologies**
   - Language (Python)
   - Packages (SpaCy, Bert, Tensorflow, Keras)
   - Jupyter Notebook
4. **Designs**
   - Component Architecture
5. **Standards and best practices**
   - Version controlling (git and GitLab)
   - Project management (MS Planner and MS Teams)
   - Python coding standards and Adding proper comments





Risk Mitigation

Expected Progress – IT19005218

Remaining Tasks

1. Rest of the implement the last model

   BERT
2. Evaluate the models
3. Testing
4. App development and integration
5. Deployment

References

[1] C. Parada, M. Dredze, and F. Jelinek, “OOV Sensitive NamedEntity Recognition in Speech.” in Proceedings of INTERSPEECH ’11, Florence, Italy, 2011, pp. 2085–2088.

[2] M. Pourakbari-Kasmaei, M. Lehtonen, J. Contreras, and J. R. S. Mantovani, ‘‘Carbon footprint management: A pathway toward smart emission abatement,’’ IEEE Trans. Ind. Informat., vol. 16, no. 2, pp. 935–948, Feb. 2020, doi: 10.1109/TII.2019.2922394.

[3] D. Nadeau and S. Sekine, “A survey of named entity recognition and classification,” Linguisticae Investigationes, vol. 30, pp. 3–26, January 2007.

[4] H. Hashim *et al.*, “An Integrated Carbon Accounting and Mitigation Framework for Greening the Industry,” *Energy Procedia*, vol. 75, pp. 2993–2998, Aug. 2015, doi: 10.1016/J.EGYPRO.2015.07.609.

Carbonis
Manage Now For Future

Component 2

• • •

Emission Factor Retrieval for Emission Calculation using Information Retrieval and Word Embeddings

Sathees P.
IT19052748
Data Science

Research Questions

Questions

1. How do calculate emissions for the emission activities efficiently and accurately?

   Emission = Consumption * Emission Factor [1], [2]

   - How to make selecting the right emission factor practical?

     - Search and rank emission factors

   - How to rank the frequently used emission factors customized for users?

     - Re-rank using personalization

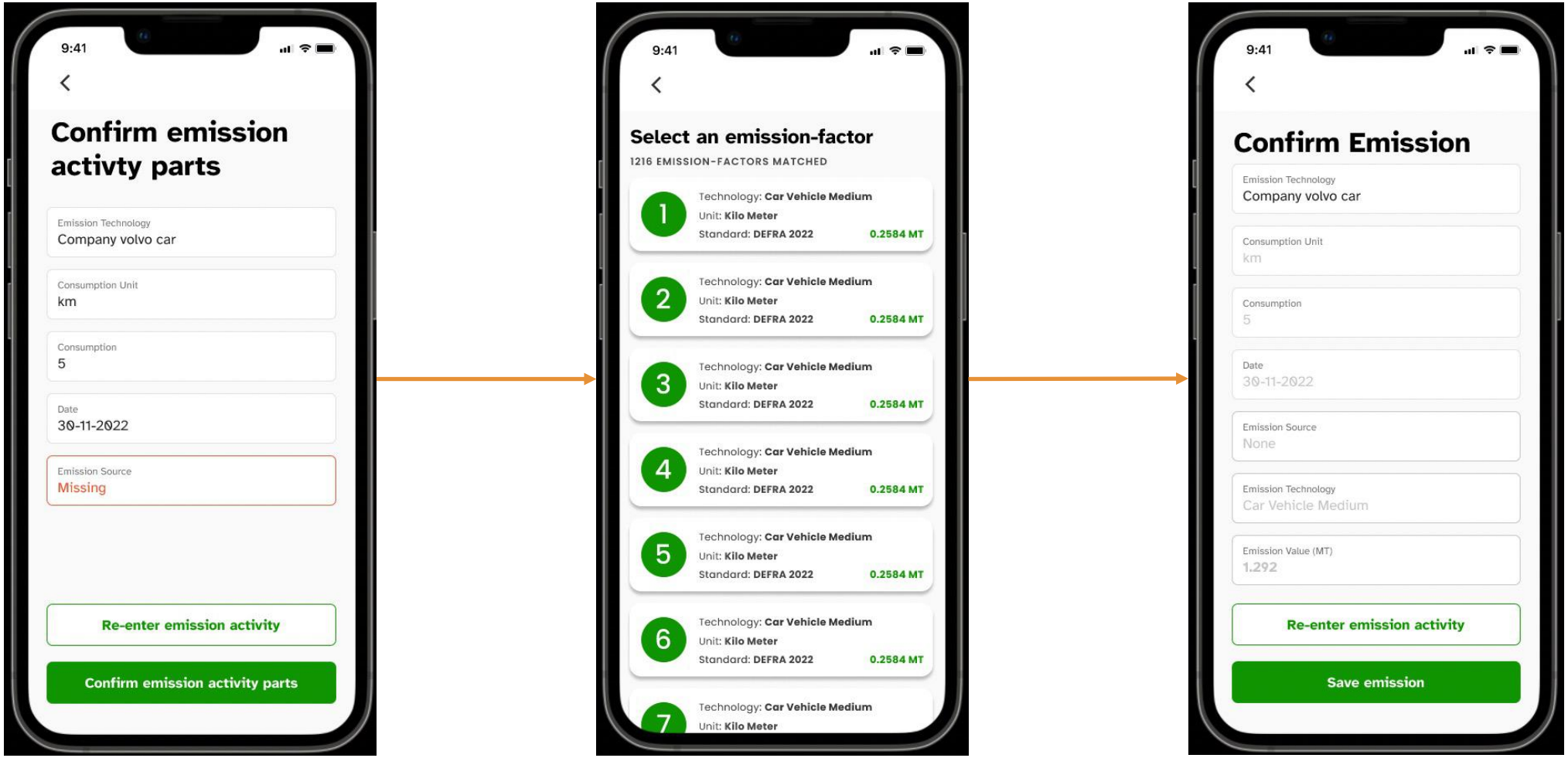
Objectives

Main Objective

Search emission factors and provide **ranked results** for the emission details gathered.

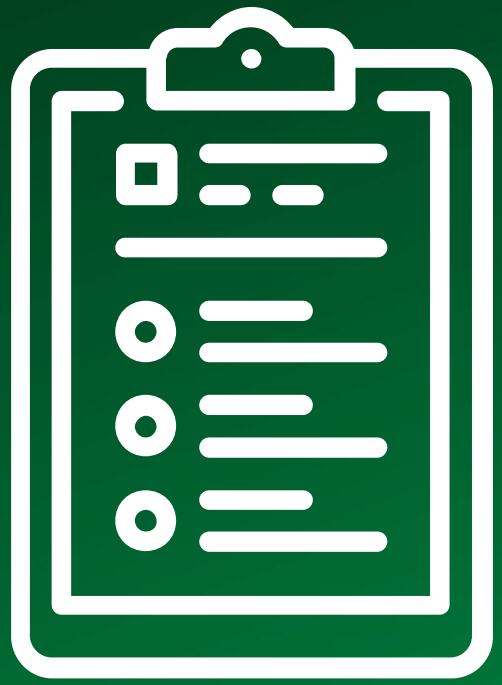
Sub Objectives

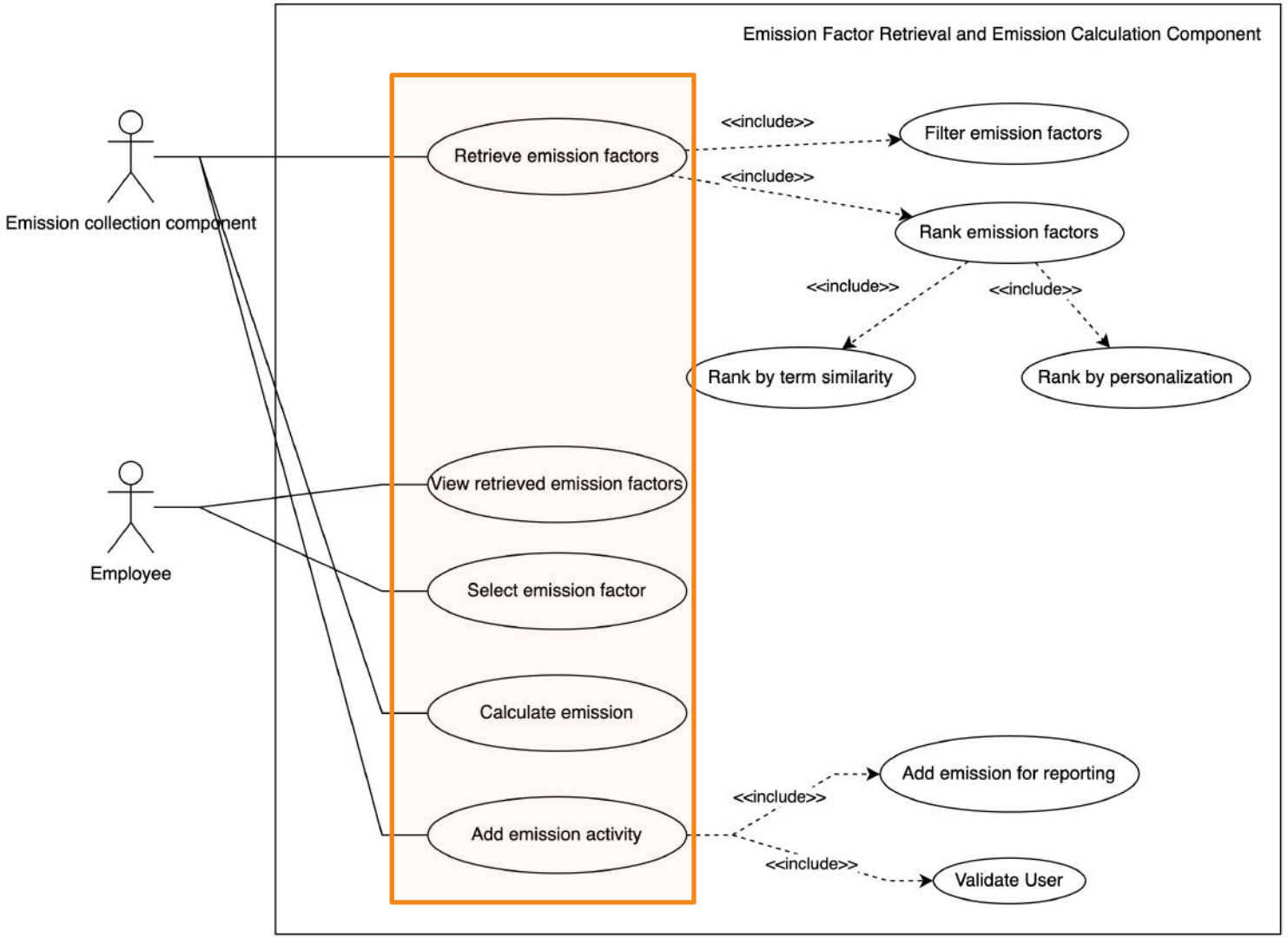
- Rank emission factors based on **term similarity** [3 - 6]
- Re-rank emission factors based on **personalization**
- Calculate emissions



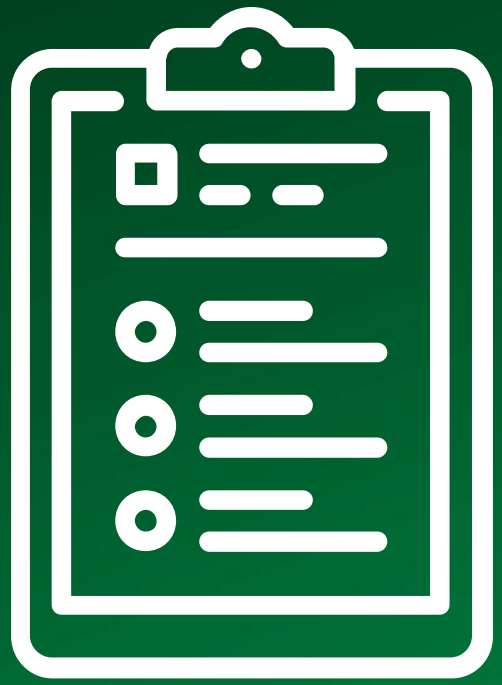
Sample UI Flow

Functional Requirements



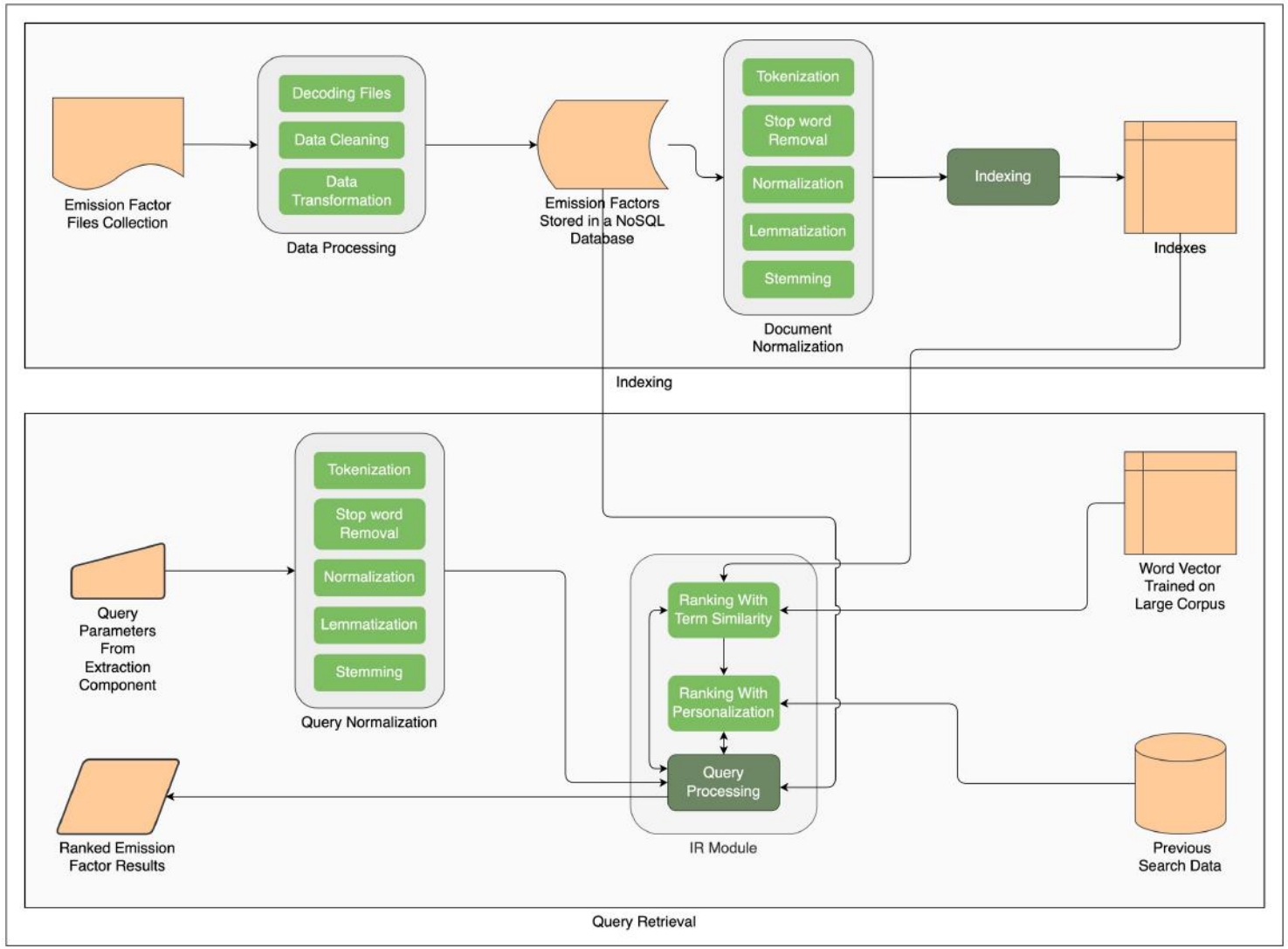


Non-Functional Requirements



Non-Functional Requirements

- Speed
- Scalability
- Ease of Use



Component 2 Architecture

Current Progress – IT19052748

Completed Tasks

1. Data collection

   From 5 standards (DEFRA, CRIS, EPA, IPCC, and NGA), 8 years (2014 - 2021), and 4 file formats (Xls, Xlsx, Pdf, Docx)

2. Component designing

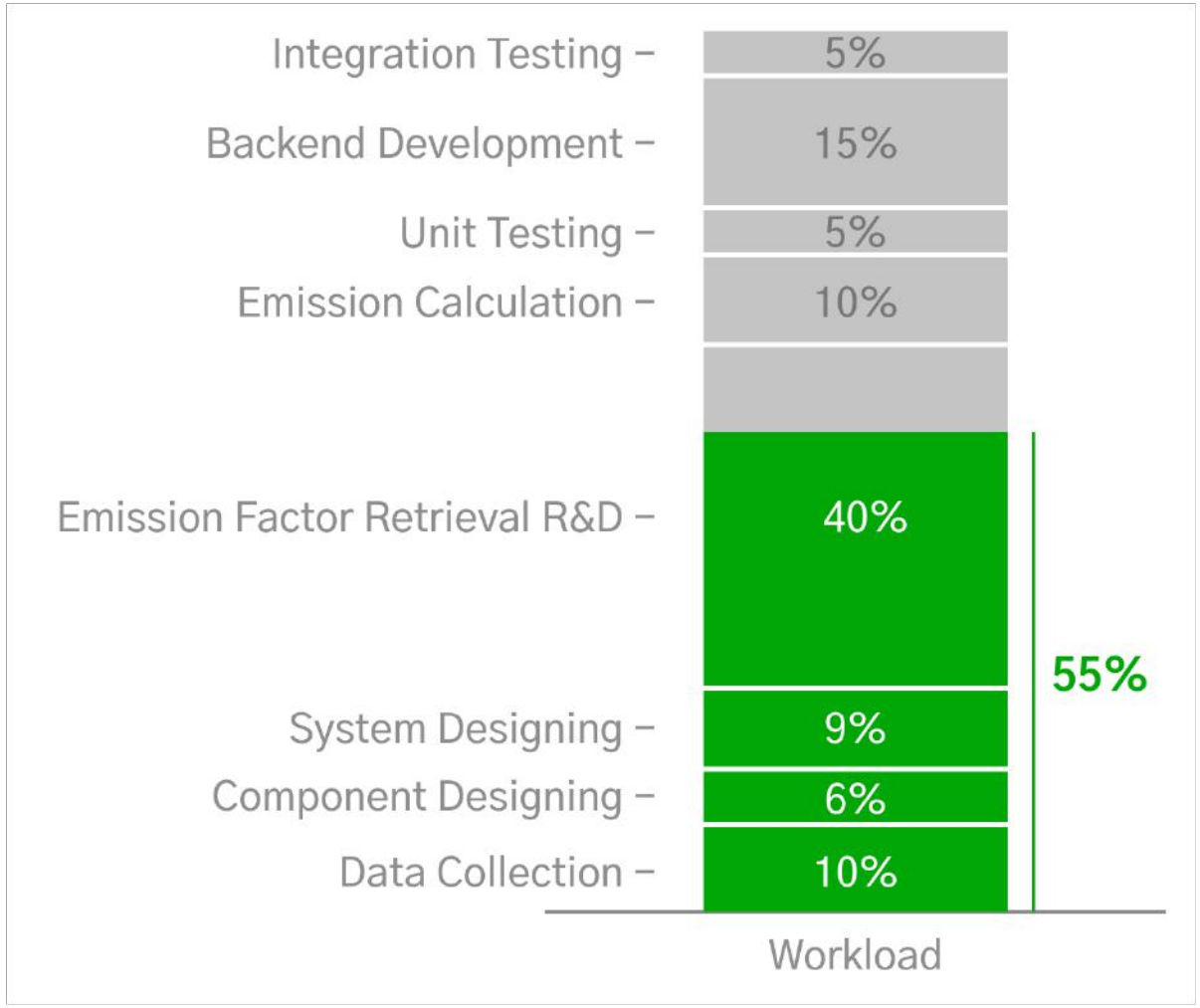
   Use case diagram, and component architecture

3. System designing

   Database design, data warehouse design, API design, and cloud architecture

4. Emission factor retrieval R&D

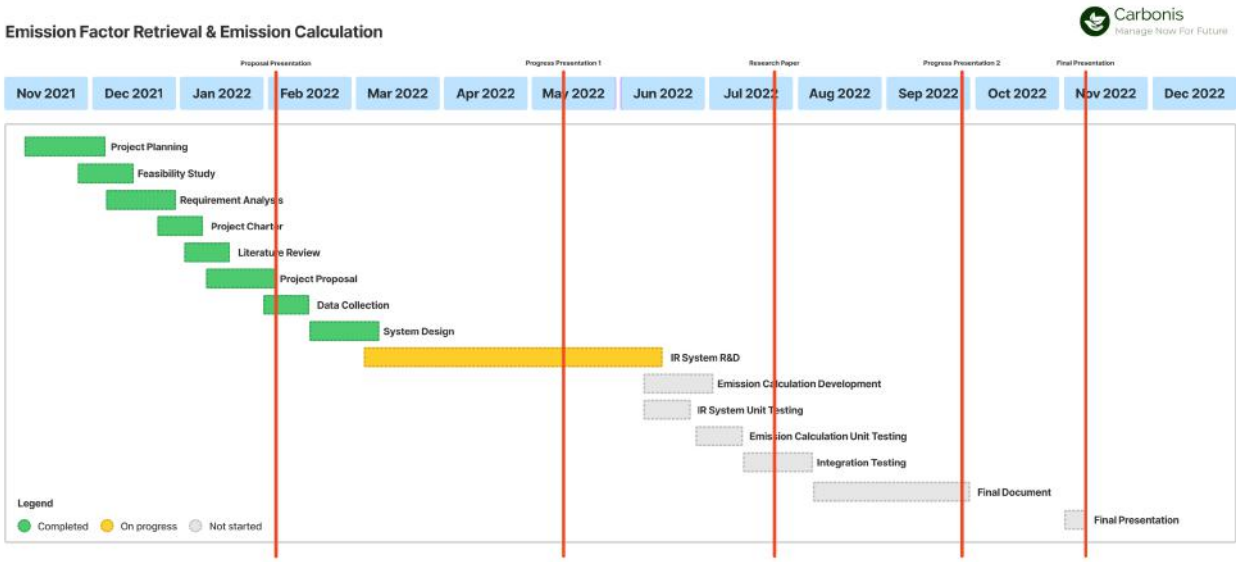
   - Data preparation
   - Common document storage
   - Document normalization
   - Indexing and index storage
   - Query normalization
   - Ranking with similarity

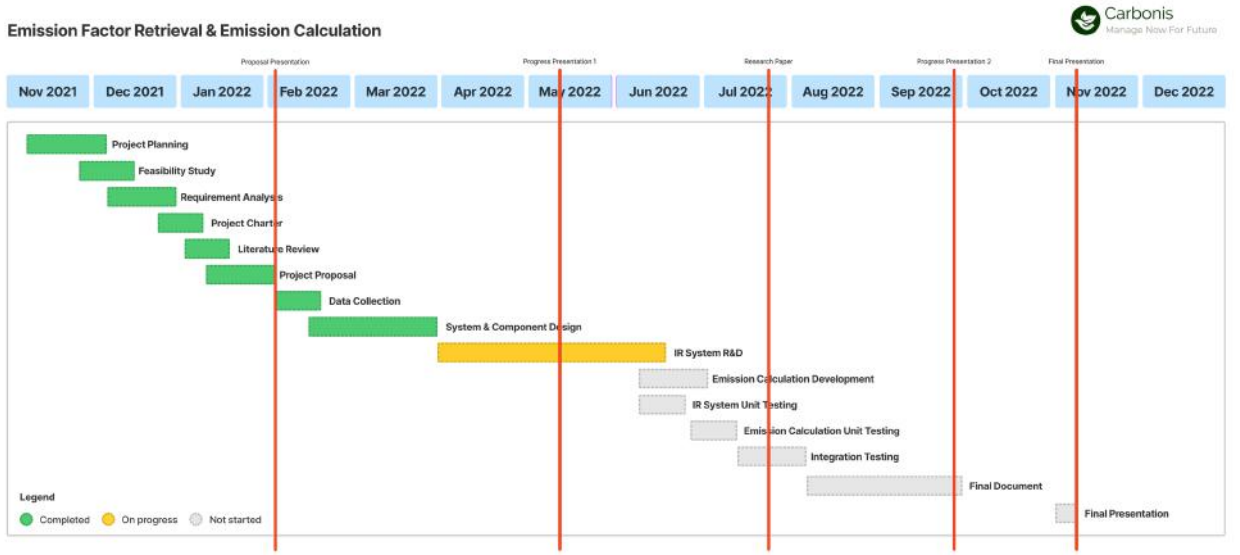


Component Workload Completion

Progress Demo (55%)

- Proof of concept
- Key pillars of the domain (Data Science)
  - Data Preparation
  - Natural Language Processing (NLP)
  - Information Retrieval (IR)
  - Word Embedding [7]
  - Databases, and Data warehousing
- Technologies
  - Language (Python)
  - Packages (Jupyter, NumPy, Pandas, Matplotlib, seaborn, NLTK, re, scikit-learn, PyMongo, Joblib, and Gensim), Database (MongoDB Atlas)
  - IDE (PyCharm)
- Designs
  - Component Architecture, Database design, Data warehouse design, Cloud design, and API design
- Standards and best practices
  - **Version controlling** (git and GitLab), **Project management** (MS Planner and MS Teams), **Separating credentials from VCS** (.env), **Python coding standards** (PEP8, and SonarLint), and Adding **proper comments**





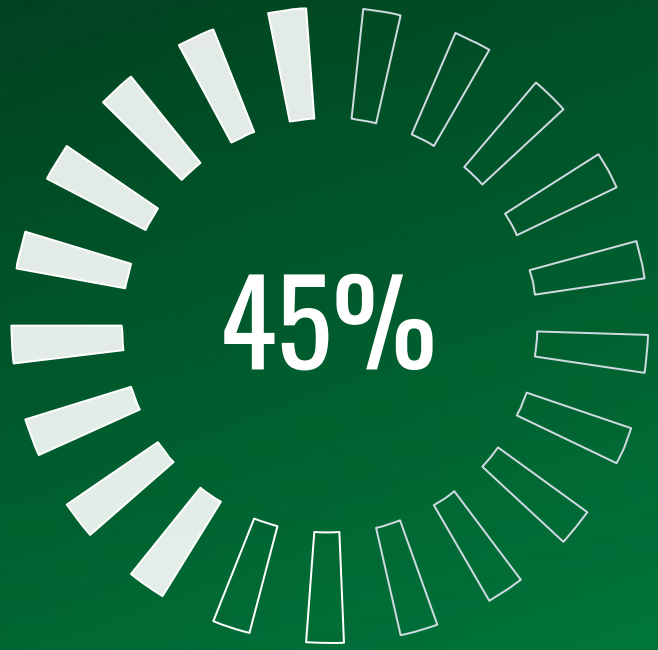
Risk mitigations

- Delay from the design phase

-> Increased the speed of the R&D phase

Updated Gantt Chart and Risk Mitigation

Expected Progress – IT19052748



Remaining Tasks

1. Emission factor retrieval R&D
   - Re-rank with personalization
   - Ranking evaluation
2. Emission calculation
3. Unit testing
4. Backend development
5. Integration testing

References

[1] "SRI LANKA UPDATED NATIONALLY DETERMINED CONTRIBUTIONS", 2021.

[2] Jayathunga, R. D. S., & Dulani, M. H. N. K. T. (2016). *A GUIDE for CARBON FOOTPRINT ASSESSMENT CLIMATE CHANGE SECRETARIAT MINISTRY OF MAHAWELI DEVELOPMENT AND ENVIRONMENT The Climate Change Secretariat Ministry of Mahaweli Development and Environment*.

[3] Ganguly, D., Roy, D., Mitra, M., & Jones, G. J. F. (2015). A word embedding based generalized language model for information retrieval. *SIGIR 2015 - Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 795–798. https://doi.org/10.1145/2766462.2767780

[4] Zamani, H., & Bruce Croft, W. (2017). Relevance-basedword embedding. *SIGIR 2017 - Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 505–514. https://doi.org/10.1145/3077136.3080831

[5] Yang, X., Lo, D., Xia, X., Bao, L., & Sun, J. (2016). Combining Word Embedding with Information Retrieval to Recommend Similar Bug Reports. *Proceedings - International Symposium on Software Reliability Engineering, ISSRE*, 127–137. https://doi.org/10.1109/ISSRE.2016.33

[6] Hu, D., Chen, M., Wang, T., Chang, J., Yin, G., Yu, Y., & Zhang, Y. (2018). Recommending Similar Bug Reports: A Novel Approach Using Document Embedding Model. *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*, *2018-December*, 725–726. https://doi.org/10.1109/APSEC.2018.00108

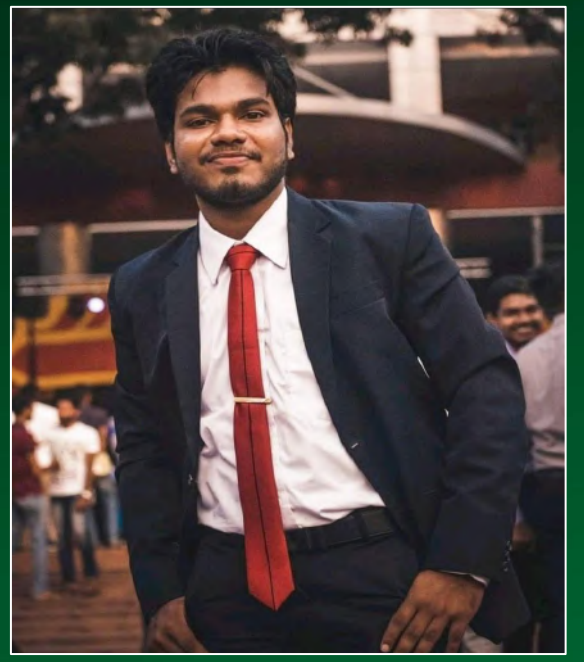
[7] J. Pennington, R. Socher, C. D. Manning, 'GloVe: Global Vectors for Word Representation', Empirical Methods in Natural Language Processing (EMNLP), 2014. 1532–1543.

Carbonis
Manage Now For Future

Component 3

• • •

Unit Verification using Text Classification And Unit Conversion



Vishakanan S.
IT19001562
Data Science

Research Questions

Questions

1. How can we make sure the units are matching in the calculation?

Solution

- Verify and convert before calculating

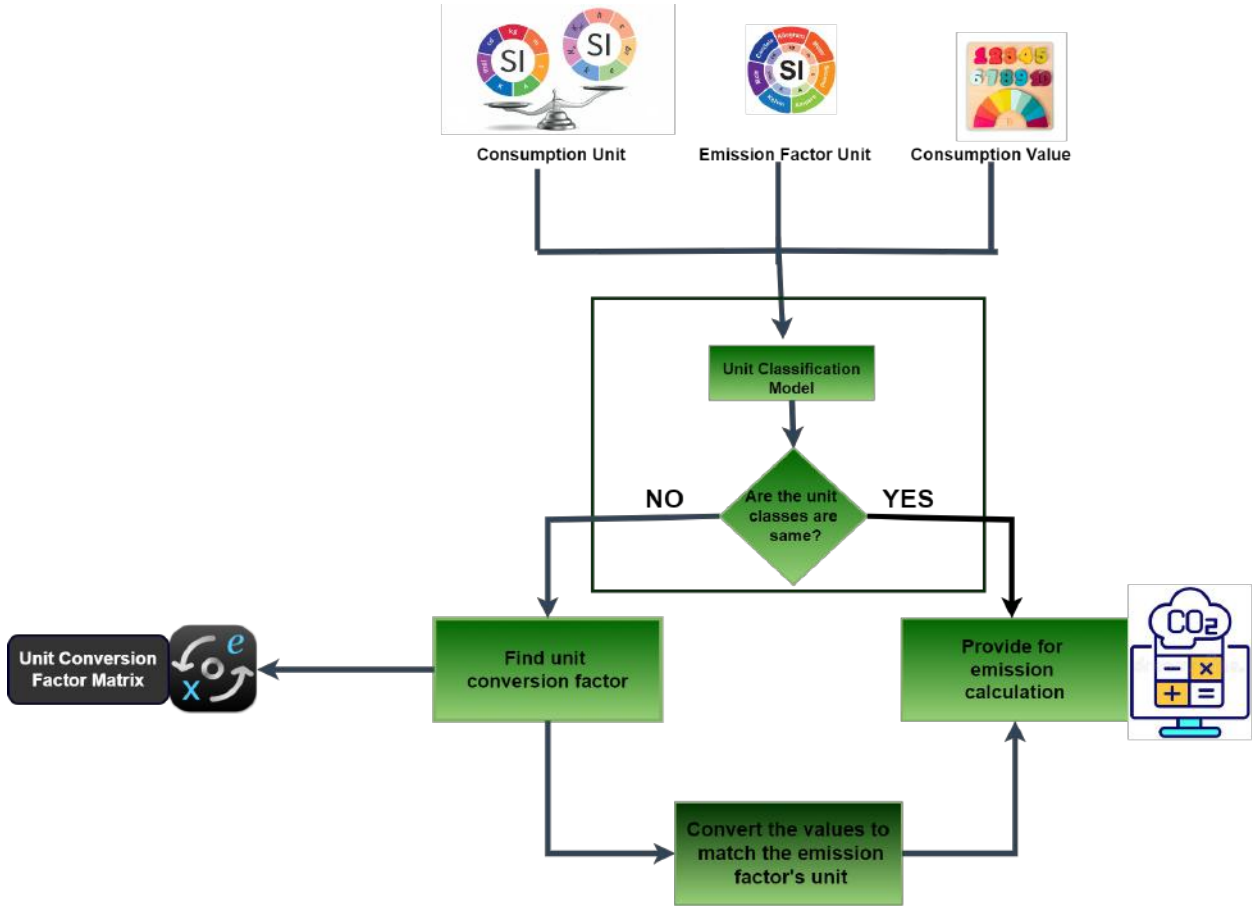
Objectives

Main Objectives

- Verify and convert consumption values, units provided to match the emission factor units.

Specific Objectives

- Unit verification using text classification.
- Unit conversion for non-matching units.

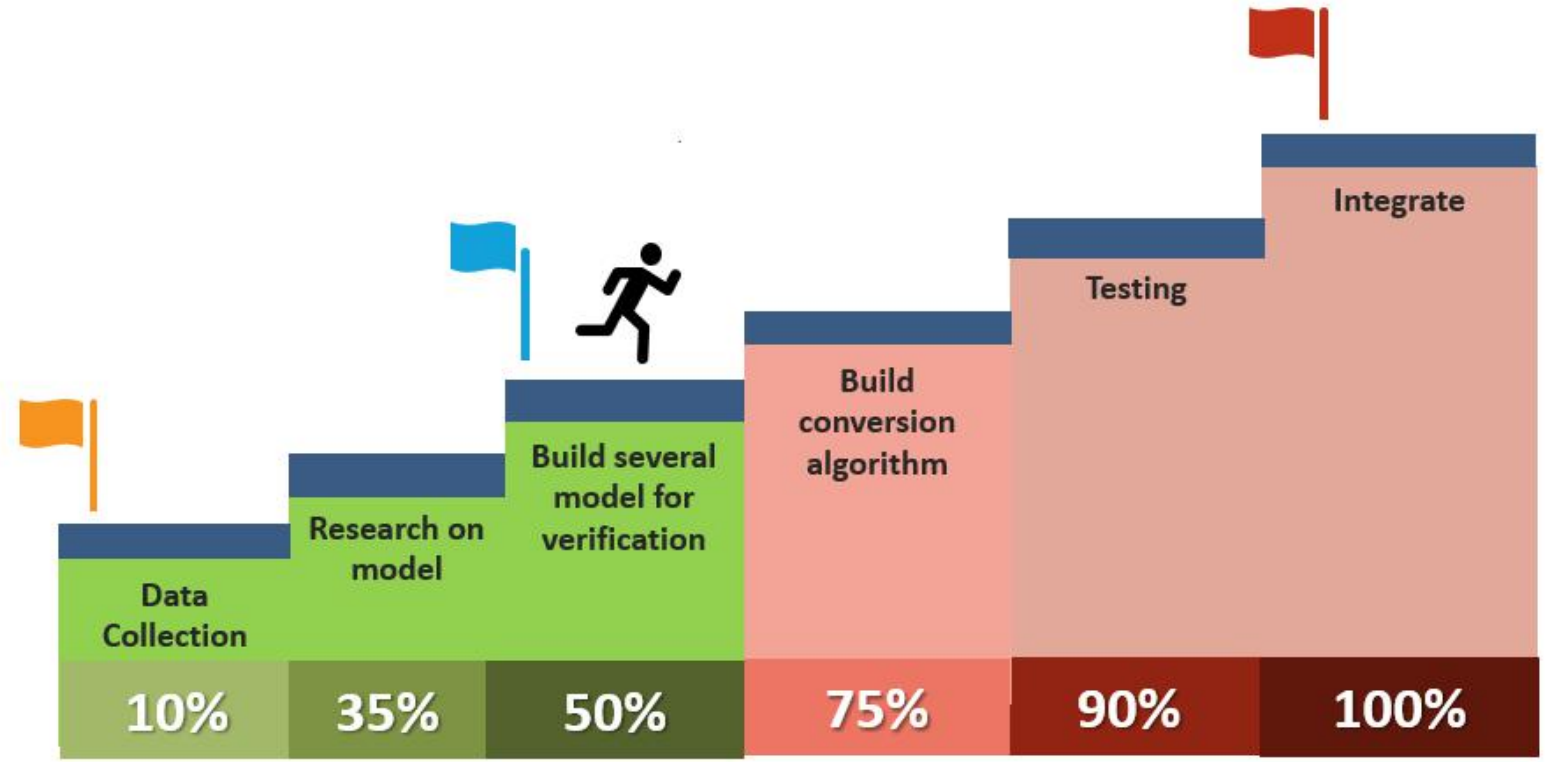


Component 3 Architecture

Current Progress – IT19001562

Completed Tasks

1. Data collection
2. Research on model selection
3. Data preprocessing
4. Text classification models implementation
5. Model comparing



Objectives and Project Completion

Progress Demo (50%)

Proof of concept

- Technologies
  - Language (Python)
  - Hugging face models
  - Transformers
  - Regax algorithm

- Standards
  - Version controlling (git and GitLab)
  - Project management (MS Planner and MS Teams),

Risk Mitigation

Expected Progress – IT19001562

Remaining Tasks

1. Build models and compare the models
2. Build unit conversion algorithm.
3. Backend development.
4. Testing
5. Integration

References

[1] Guidance on how to measure and report your greenhouse gas emissions. Department for Environment, Food and Rural Affairs., 2009, pp. 20-22.

[2] Carbon Footprint of an Organization: a Tool for Monitoring Impacts on Global Warming, Department of Agricultural Engineering, Faculty of Agriculture, University of Ruhuna, Mapalana, Kamburupitiya, 81100, Sri Lanka. 2017

[3] Estimation of renal function in the intensive care unit: the covert concepts brought to light Sham Sunder, Rajesh Jayaraman*, Himanshu Sekhar Mahapatra, Satyanand Sathi, Venkata Ramanan, Prabhu Kanchi,Anurag Gupta, Sunil Kumar Daksh and Pranit Ram, 2014

[4] Android based Conversion and Estimation Application, March 2016

[5] Measurement Context Extraction from Text: Discovering Opportunities and Gaps in Earth Science, Kyle Hundman1, Chris A. Ma□mann1,2

[6] Measurement Context Extraction from Text: Discovering Opportunities and Gaps in Earth Science, Kyle Hundman1, Chris A. Ma□mann1,2

[7] Automated Detection of Measurements and Their Descriptors in Radiology Reports Using a Hybrid Natural Language Processing Algorithm

[8] How to Extract Unit of Measure in Scientific Documents? , KDIR 2013

[9] Natural Language Processing Techniques for Extracting and Categorizing Finding Measurements in Narrative Radiology Reports, 2015

Carbonis
Manage Now For Future

Component 4

• • •

Emission Optimization using Linear Programming

Vithursan M.
IT19033174
Software Engineering

Research Questions

Questions

- How to reduce the emission?
- How to find the threshold values for each emission sources?
- How to maintain the carbon emission level without exceeding the limit?

Objectives

Main Objectives

Identify the optimum solution for the given emission source constraints using **Optimization Algorithms** and sent alert about any violations of the optimal solution.

Specific Objectives

1. Implementing a custom emission optimization module.

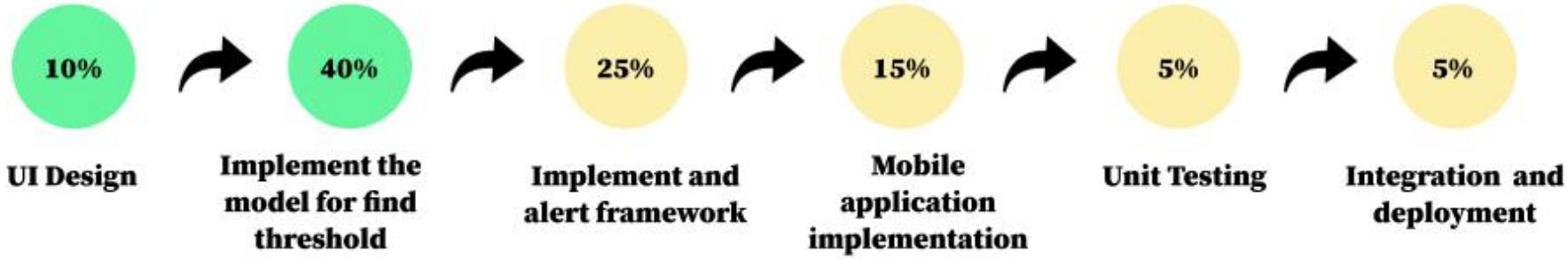
2. Creating an alert framework to provide alerts about the breaches of the thresholds.

3. Implement a mobile application using React Native.

Current Progress – IT19033174

Completed Tasks

1. Mobile application UI wireframe
2. High fidelity prototype
3. Data collection
4. Optimization model to find threshold
5. Mobile UI implementation



Objectives and Project Completion

Progress Demo (50%)

Expected Progress – IT19033174

Remaining Tasks

1. Alert framework for any violation of threshold
2. Rest of mobile application implementation
3. Integration with other components
4. Testing
5. Deployment

References

[1] William F Lamb, Thomas Wiedmann, Julia Pongratz , Robbie Andrew , Monica Crippa, Jos G J Olivier, Dominik Wiedenhofer "A review of trends and drivers of greenhouse gas emissions by sector from 1990 to 2018" *Environmental Research Letters , Volume 16, Number 7 ,* Published on 29 June 2021 ,Published by: IOP Publishing Ltd available at: https://iopscience.iop.org/article/10.1088/1748-9326/abee4e

[2] M. Roelfsema et al., "Taking stock of national climate policies to evaluate implementation of the Paris Agreement", *Nature Communications*, vol. 11, no. 1, 2020. Available at: https://www.nature.com/articles/s41467-020-15414-6?fbclid=IwAR1drArL9ReoJl2zgqjmdxJNoBsM4zRJna-JHIGWkzTka7d4NB4fdz0nCrE.

[3] B. Tranberg, O. Corradi, B. Lajoie, T. Gibon, I. Staffell and G. Andresen, "Real-time carbon accounting method for the European electricity markets", *Energy Strategy Reviews*, vol. 26, p. 100367, 2019. Available: https://www.sciencedirect.com/science/article/pii/S2211467X19300549.

[4] Kazi Mostafa, Innchyn Her, "Stabilization wedges as a tool of engineering optimization, with an example of CO2 emission control", vol. 1, 2010. available at: https/ieeexplore.ieee.org/document/5533732

[5] E. T. Lau, Q. Yang, G. A. Taylor, A. B. Forbes, P. Wright, V. N. Livina, "Optimization of carbon emissions in smart grids" , 2014, ISBN:978-1-4799-6557-1, . available at: https/ieeexplore.ieee.org/document/6934796

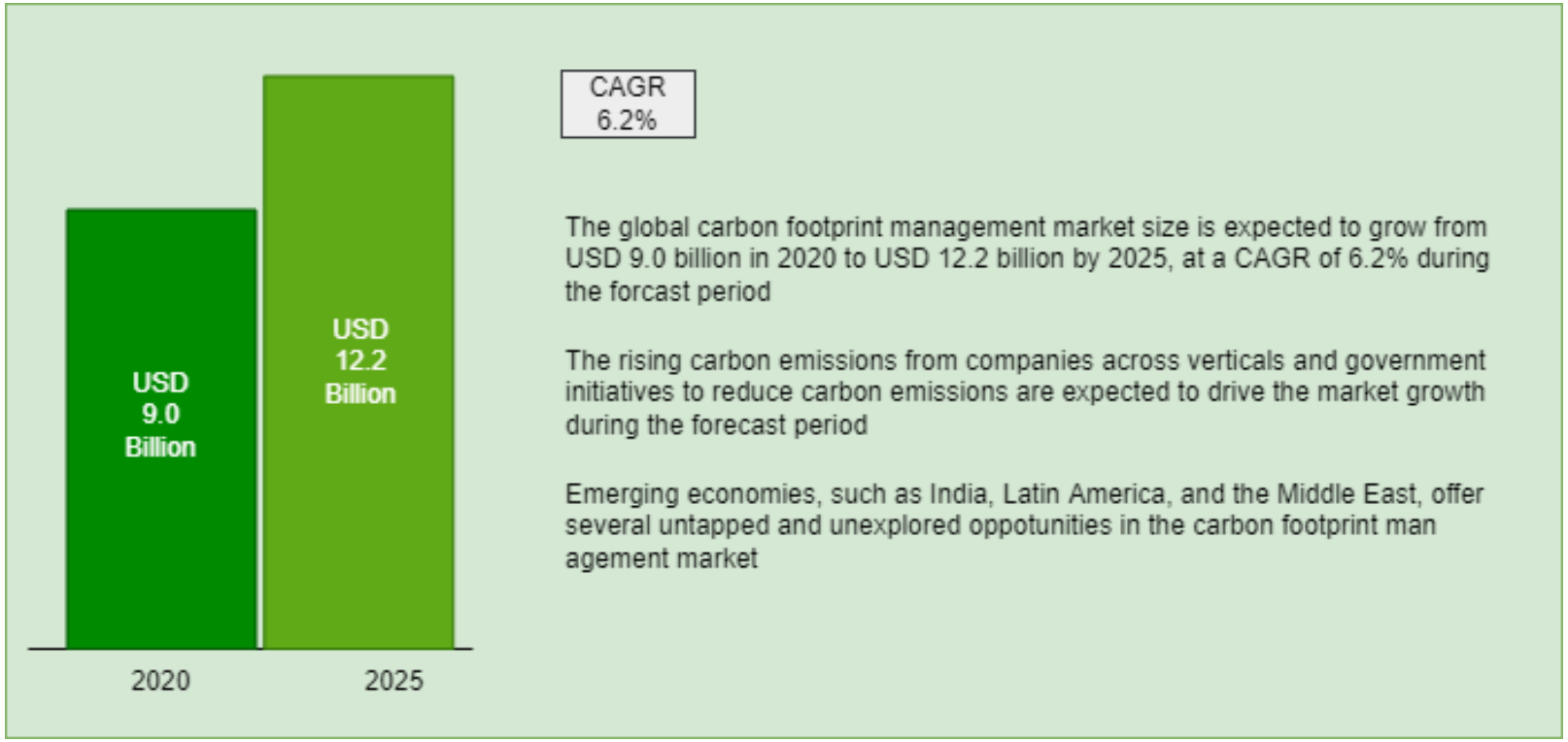
[6] Kailong Zhou, Xin Chen, Weihua Cao, "Optimization Method for Carbon Efficiency in the Green Manufacturing of Sinter Ore and its Application", 2018, ISBN : 1934-1768, available at: https/ieeexplore.ieee.org/document/8483152

Commercialization

Commercialization Methodology

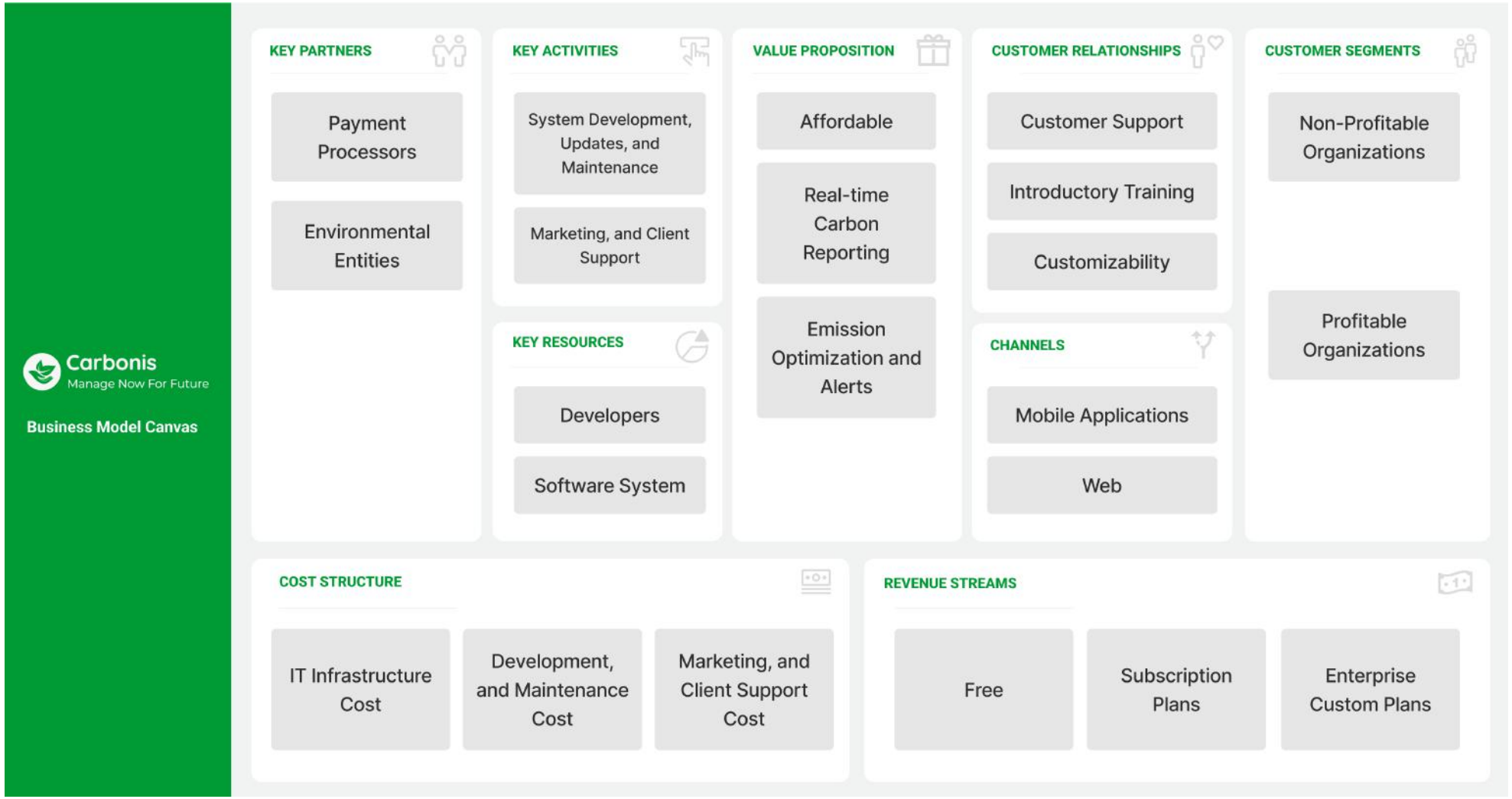
1. Market Analysis
2. Business Modeling
3. Business Plans

Attractive opportunities in the carbon footprint management market

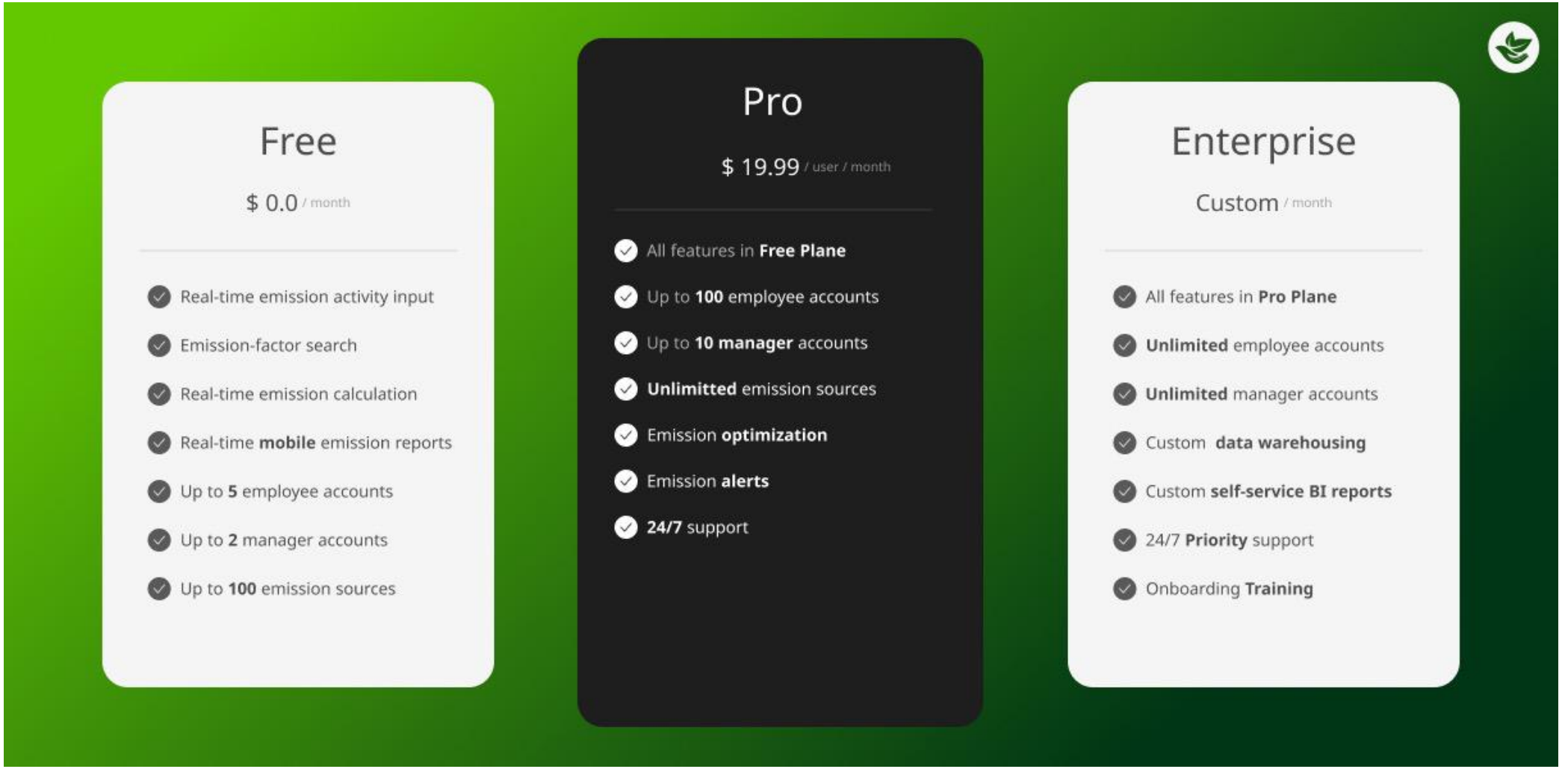


Source : Secondary research, markets and market analysis

1. Market Analysis



2. Business Model Canvas



3. Business Plans

Thanks!

Do you have any questions?