Semester Long Internship Report

on

Unified Emissions Intelligence: A Comprehensive Framework for Scope 3 Data Systems, Event Reporting, and Brightway2 Modeling

Submitted by

Ammulu Uppalapati

Rajiv Gandhi University of Knowledge Technologies

(Computer Science and Engineering)

Under the guidance of

Prof. Kannan M. Moudgalya

Chemical Engineering Department, IIT Bombay

Mentors

Mr. Shubham Sonkusare

Mr. Nikhil Sharma

Under Supervision of

Mr. Sumanto Kar

(Asst. Project Manager)

July 9, 2025

Declaration

I hereby declare that this internship report titled **“Scope 3 Emission Data Collection, Analysis, and Dashboard Development Using Open Source Tools”** is a genuine and original work submitted in partial fulfillment of the requirements for the FOSSEE Semester-Long Internship, conducted by the Free/Libre and Open Source Software for Education (FOSSEE) team at IIT Bombay.

This report has not been submitted to any other institution or university for the award of any degree, diploma, or academic credit. Any literature, data, or information from other sources has been duly acknowledged and appropriately cited in the report.

I have abided by the academic integrity standards and ethical practices throughout the course of this internship and while preparing this report.

Name: Ammulu Uppalapati
Registration Number: [Insert your reg. no. if needed]
Program: B.Tech, Computer Science and Engineering
Institution: Rajiv Gandhi University of Knowledge Technologies
Internship Duration: Febraury 2025 – July 2025

Acknowledgement

I would like to take this opportunity to express my sincere gratitude to the **FOSSEE team at IIT Bombay** for granting me the privilege of participating in the Semester-Long Internship Program. This internship has been a valuable experience that helped me strengthen my technical knowledge and explore the intersection of open-source technology and sustainability.

I am deeply thankful to my mentors — **Mr. Shubham Sonkusare**, **Mr. Sumanto kar**, and **Mr. Nikhil Sharma** — for their unwavering support, timely feedback, and continuous encouragement throughout the internship period. Their guidance was instrumental in shaping my project work and enhancing my learning.

I would also like to thank my home institution, **Rajiv Gandhi University of Knowledge Technologies**, for their academic and logistical support that enabled me to take part in this internship.

Lastly, I am grateful to my peers, colleagues, and the wider FOSSEE community, whose collaborative spirit and dedication inspired me to contribute effectively to this impactful initiative.

Abstract

This report outlines my semester-long internship experience with the FOSSEE (Free/Libre and Open Source Software for Education) project at IIT Bombay. The primary objective of the internship was to contribute towards environmental sustainability by analyzing greenhouse gas (GHG) emissions, with a strong focus on **Scope 3 emissions**.

The internship commenced with a screening task, where I was required to collect and analyze Scope 3 emission data using Python. This task helped build a strong foundation in understanding emission categories, data sources, and the computational methods required for processing and visualization.

After successful selection, I worked on building a web-based **interactive emissions dashboard** for a sample event. This dashboard was developed using **Python**, **Stream lit**, and **SQLite**, and was capable of calculating and displaying **Scope 1, Scope 2, and Scope 3 emissions**. To make the system dynamic and user-driven, we developed **custom forms** to collect information such as attendees' mode of transport, travel distance, and food preferences—key contributors to event-based emissions.

In the later stages, I was assigned a specific category under Scope 3 emissions for in-depth data collection and structuring. The work involved identifying reliable sources, compiling emission factors, and preparing the data for integration into the system. Additionally, I was introduced to Brightway2, a Python-based Life Cycle Assessment (LCA) tool, which I began exploring for modeling environmental impacts at a more granular level.

This internship provided valuable insights into sustainable development, environmental data analysis, and full-stack Python development. It offered a unique combination of domain knowledge and technical skill development, empowering me to contribute meaningfully to real-world sustainability projects using open-source tools.

Table of Contents

1. Introduction

1.1 About FOSSEE

The **FOSSEE (Free/Libre and Open Source Software for Education)** project is an initiative by the Indian Institute of Technology Bombay (IIT Bombay), supported by the Ministry of Education under the National Mission on Education through ICT (NMEICT). Its mission is to promote the use of open-source software in education and research. FOSSEE provides various internship programs for students to contribute to impactful, real-world open-source projects in different domains such as Python, Scilab, OpenFOAM, Sustainable Computing, and more.

1.2 Importance of Open-Source Tools

Open-source tools play a vital role in democratizing access to technology. They are cost-effective, highly customizable, and supported by active communities. In academic and sustainability projects, open-source platforms enable transparency, reproducibility, and flexibility. During this internship, tools like **Python**, **Streamlit**, **SQLite**, and **Brightway2** were utilized extensively to build and visualize emission tracking systems and perform life cycle assessments.

1.3 Overview of Environmental Emissions

Environmental sustainability has become an essential area of focus globally. A critical aspect of this is tracking and reducing **greenhouse gas (GHG) emissions**, which contribute significantly to global warming. These emissions are broadly categorized into three types, defined by the **Greenhouse Gas Protocol**:

- **Scope 1**: Direct emissions from owned or controlled sources
- **Scope 2**: Indirect emissions from the generation of purchased energy
- **Scope 3**: All other indirect emissions occurring in the value chain

Scope 3 emissions, which are the most diverse and complex, include activities such as transportation, procurement, employee commuting, food consumption, and waste.

1.4 Understanding Scope 1, 2, and 3 Emissions

Scope	Description	Examples
Scope 1	Direct emissions from sources controlled by the organization	Fuel combustion, company-owned vehicles
Scope 2	Indirect emissions from purchased electricity or energy	Grid electricity, heating
Scope 3	Indirect emissions across the value chain	Business travel, food choices, transport, waste management

This internship focused heavily on **Scope 3 emissions**, especially from events - including emissions from food consumption and transportation by attendees. Understanding and calculating these emissions is vital for any organization aiming to become more environmentally responsible.

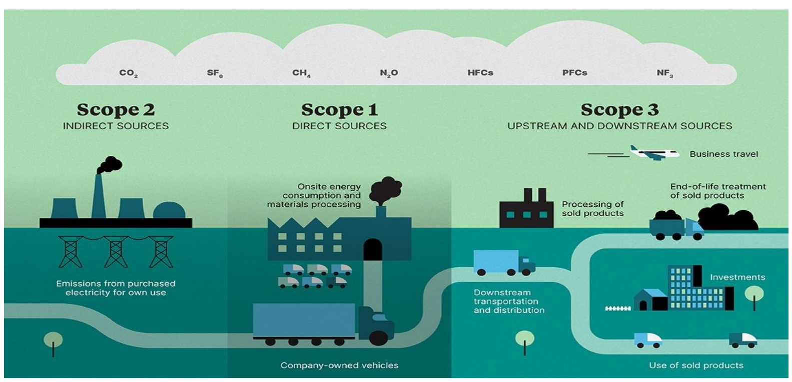

Figure 1: GHG emissions by Scope 1, 2, and 3.

2. Internship Overview

2.1 Internship Title and Domain

The title of my internship was:

"Scope 3 Emission Data Collection, Analysis, and Dashboard Development using Open Source Tools"

This internship was offered under the **Sustainable Computing domain** of the FOSSEE project at **IIT Bombay**. It focused on the analysis and visualization of greenhouse gas emissions, particularly **Scope 3 emissions**, using **Python** and open-source libraries. The project also included the design and development of a real-time **web dashboard** using **Streamlit**, alongside work related to **Life Cycle Assessment (LCA)** using **Brightway2**.

2.2 Internship Duration

- **Start Date:** February 2025
- **End Date:** July 2025
- **Mode:** Remote (Online)
- **Type:** Semester-long Internship

2.3 Selection Process

The selection process involved the following stages:

- **Application Submission:** Interested candidates had to apply through the FOSSEE internship portal.
- **Screening Task:** Applicants were required to complete a screening task focused on **Scope 3 emissions data collection and analysis using Python**.
- **Review and Evaluation:** The FOSSEE team evaluated submissions based on the quality of code, data handling, documentation, and understanding of emission concepts.
- **Final Selection:** Selected students were notified via email and onboarded to begin the internship.

2.4 Objectives of the Internship

The main objectives of the internship were:

- To understand the structure and importance of Scope 1, 2, and 3 emissions.
- To perform structured **data collection** for Scope 3 emission categories.
- To analyze emissions using **Python** and visualize them in a user-friendly format.
- To develop a **Streamlit-based web application** for emission tracking during events.
- To explore **Brightway2**, an open-source Life Cycle Assessment (LCA) tool.
- To contribute toward open-source solutions for sustainability.

2.5 Tools and Technologies Used

Tool / Technology	Purpose
Python	Data collection, processing, and scripting
Pandas, NumPy	Data analysis and transformation
Streamlit	Web dashboard development
SQLite	Local database for form data storage
Matplotlib / Plotly	Data visualization
Brightway2	Life Cycle Assessment modeling
Git & GitHub	Version control and collaboration
Markdown / LaTeX	Documentation

3. Screening Task

3.1 Objective

The screening task served as a prerequisite for selection into the semester-long FOSSEE internship. The objective was to assess candidates' ability to:

- Understand the concept of **Scope 3 emissions**
- **Collect reliable data** for one or more Scope 3 categories
- Perform **data analysis and cleaning** using Python
- Present initial insights using **code and visualizations**

This task provided foundational exposure to environmental sustainability topics and allowed interns to demonstrate their technical and analytical skills.

3.2 Task Description

The screening task focused on **Scope 3 Emissions – Data Collection and Analysis**. Scope 3 emissions, unlike Scope 1 and 2, are indirect emissions that occur throughout the value chain of an organization. The goal was to select one or more categories (such as purchased goods, waste, or transportation) and:

1. Collect **real and verifiable data** from open datasets, reports, or scientific papers.
2. Use **Python** to read, clean, and process the data.
3. Organize the data into structured formats (e.g., CSV, JSON).
4. Conduct **basic analysis** such as calculating emissions, category-wise comparison, etc.
5. Visualize data using **charts** or **tables**.

3.3 Tools and Libraries Used

Tool/Library	Purpose
Python	Main programming language
Pandas	Data manipulation and cleaning
NumPy	Array and numerical operations
Matplotlib	Data visualization (bar/pie charts)

3.4 Workflow

1. **Understanding Scope 3 Categories:**
   Researched various categories such as transportation, food services, waste generation, and employee travel.

2. **Data Collection:**
   Used open data sources such as:
   - Government databases
   - Emission factor reports
   - Published academic papers

3. **Data Cleaning:**
   - Removed duplicates and null values

- Handled inconsistent units (e.g., kg $CO_2 e, tons, etc.) Converted values into standard formats$

4. **Analysis Performed:**
   - Total emissions per category
   - Average emission per item/unit
   - Comparison charts (e.g., mode of transport vs. emissions)

5. **Visualization:**
   - Bar charts for emission distribution
   - Pie charts for percentage contribution
   - Tables for emission factor references

3.5 Outcome

- Successfully completed the screening task using open-source data and Python.
- Submitted structured data, visualizations, and well-commented Python code.
- Gained hands-on experience with environmental datasets.
- Shortlisted and selected for the semester-long internship.

4. Project Phase 1 – Event Emission Dashboard and Carbon Sequestration Study

4.1 Overview

The first phase of my internship focused on developing a system to track and analyze carbon emissions from physical events. The goal was to build a web-based dashboard that collects attendee data — including travel modes and food preferences — to calculate emissions across Scope 1, 2, and 3. I was specifically responsible for implementing the logic for Scope 1 (direct emissions) and Scope 2 (indirect energy emissions). This phase blended software development with environmental research. As the project progressed, I also explored carbon sequestration methods and compiled scientifically validated data to estimate potential offsets.

4.2 Emission Dashboard Development

4.2.1 Objective The core objective was to create an easy-to-use dashboard where participants of an event could enter their personal choices (like transport and diet), and the system would automatically calculate the carbon footprint. The dashboard also provided organizers a visual summary of the total emissions generated during the event.

4.2.2 Dashboard Visualization An essential part of this project was designing a user-friendly, interactive interface that guides the user from event registration to emissions data entry and finally to a visual representation of calculated greenhouse gas (GHG) emissions.

After data is entered, the dashboard provides instant feedback. Visualizations include:

- A real-time carbon footprint summary
- Category-wise emission breakdown (e.g., pie charts, bar graphs)
- Scope-wise contribution (Scope 1, 2, 3)
- Progress indicators showing emissions vs. reduction targets

The dashboard also includes a chatbot assistant to guide users in understanding each section, ensuring accessibility even for non-technical users.

Step 1: Register Event Users begin by registering a new event with details such as name, date, location, and guest count.

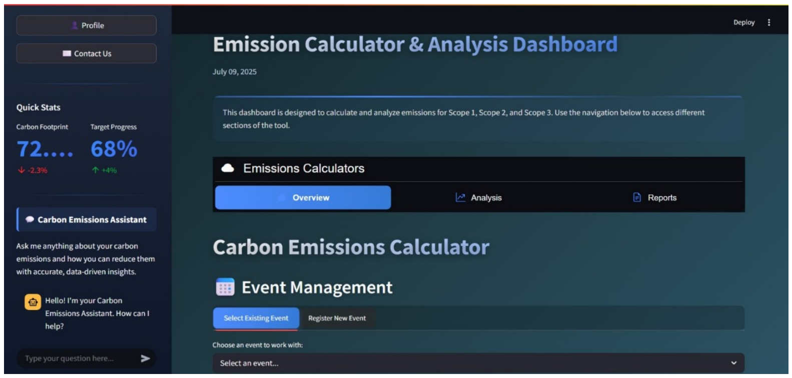

Figure 2: Home interface of the Carbon Emissions Calculator showing event management, Scope selection, and real-time carbon footprint insights.

Step 2: Scope 1 Emission Calculator

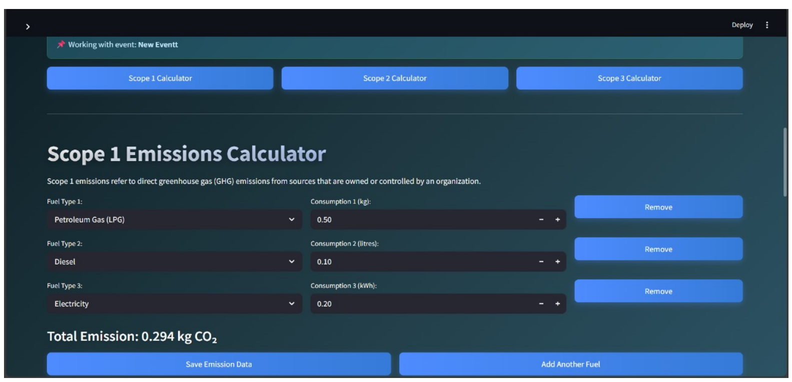

Figure 3: User interface for Scope 1 emissions calculation based on fuel consumption.

Step 3: Scope 2 Emission Calculator

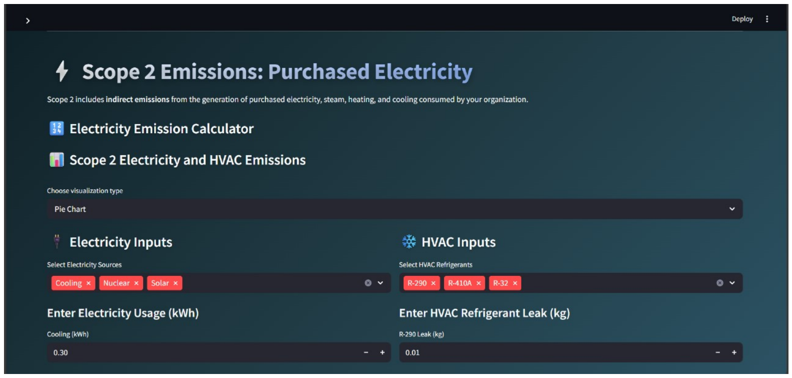

Figure 4: User interface for Scope 2 emissions calculation from electricity and HVAC sources.

Step 4: Transportation Emission calculation data collected using Form

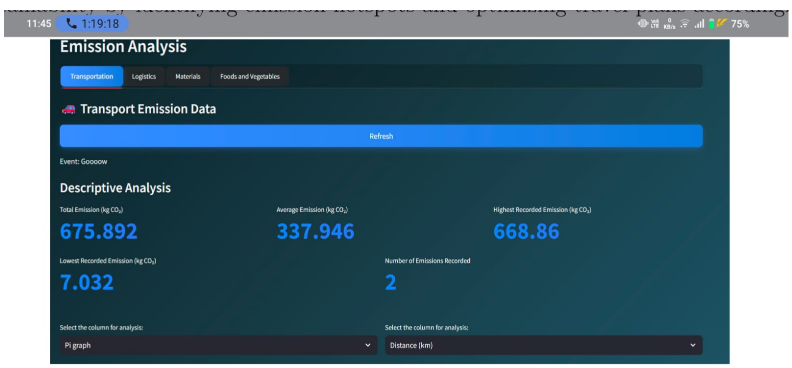

Figure 5: Scope 3 emissions calculation from Transportation

Step 5: Emission Analysis of a Event

Figure 6: Carbon Emissions Breakdown for Sample Event 'Newly'

4.2.3 My Contribution: Form Design and Dashboard Development

One of my main contributions in this phase was developing the **form system**, which formed the backbone of data collection.

I built multiple user-friendly forms using **Streamlit**:

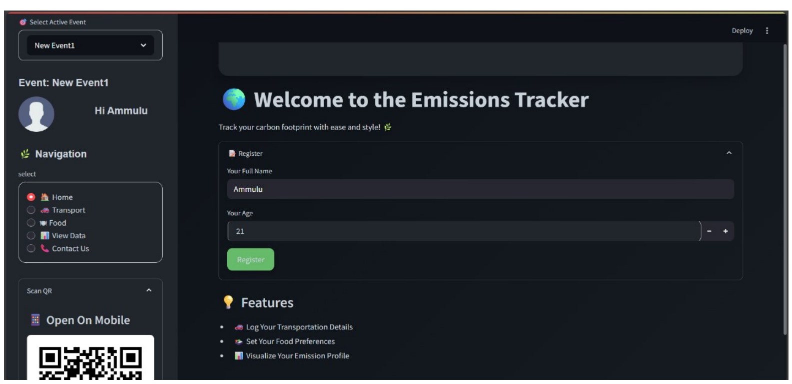

Figure 7: User registration screen of the Emissions Tracker application.

- **Registration Form** to collect basic attendee details like name and age.
- **Transport Form** to log travel details, such as distance and mode of travel.
- **Food Preference Form** to capture dietary patterns, which significantly affect emissions.

- **Contact Form** to receive feedback or queries from users.

Each form was **connected to a SQLite database** in the backend, allowing data to be saved and retrieved efficiently. I also implemented **session tracking** so each user's data stayed consistent throughout their interaction.

To make the dashboard more accessible, I even added a **QR code feature** that lets users open the app directly on their phone — useful for live events.

This was not just about coding — it was about designing an experience that was simple, responsive, and meaningful for users while ensuring that the data could be used for accurate emission tracking.

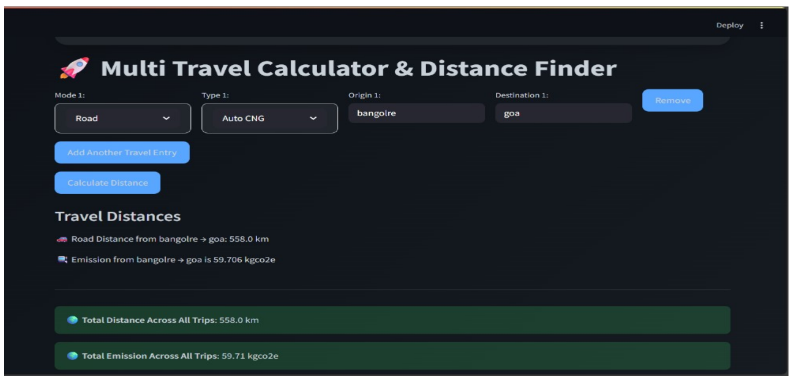

Figure 8: Travel calculation using selected mode and fuel type.

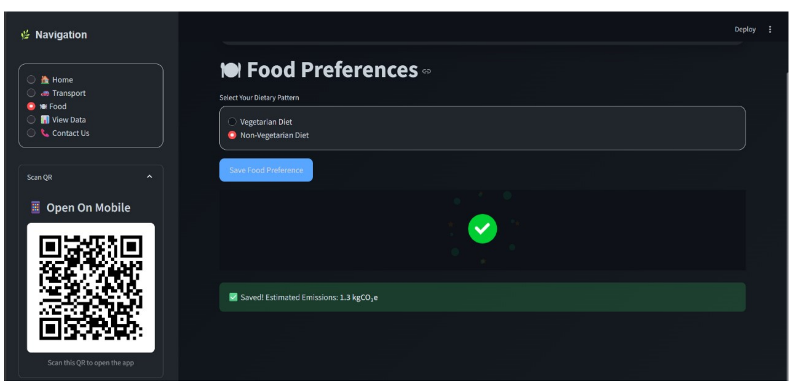

Figure 9: Emissions Calculations from Food Preferences.

4.2.4 Emission Calculation Logic Once the data was collected through the forms, it was processed using emission factors mapped to each activity. For example:

- Train travel was multiplied by 0.041 kg COe/km
- A non-vegetarian meal added 5.0 kg COe

The logic was handled in Python, and I ensured that the form values were correctly passed to these calculations.

4.2.5 Tools and Technologies Used

- **Python** for scripting and backend logic
- **Streamlit** for UI development
- **SQLite** for storing user and event data
- **Matplotlib / Plotly** for visualization
- **Qrcode / UUID / Time libraries** for session handling and access features

4.2.6 Challenges I Faced Some of the key challenges included:

- Managing real-time data flow between the form and the backend
- Preventing duplicate entries and ensuring each session was unique
- Styling Streamlit components to look clean and work well across devices

Through careful debugging and incremental testing, I was able to overcome these and deliver a stable, intuitive system.

4.2.7 Outcome By the end of this phase, I had developed a **fully functional form-driven dashboard** that could:

- Accept real-time user data
- Calculate emissions instantly
- Store and visualize results
- Scale across multiple events

This was my first time building something that combined data science with real-world environmental impact, and it was a rewarding experience.

4.3 My Research on Carbon Sequestration

After completing the dashboard, I also spent time researching **carbon sequestration** — the process of absorbing carbon dioxide from the atmosphere through natural means like forests, soil, and wetlands.

What I Explored:

- Types of sequestration: **biological** (trees, soil) vs. **geological**
- Carbon capture potential of different **tree species and ecosystems**
- How to use sequestration to **offset emissions** and move toward **net-zero**

Data I Collected: I collected comprehensive carbon sequestration data from credible sources such as the IPCC, FAO, scientific journals, and government reports. The dataset included CO absorption capacities of commonly found Indian tree species like Neem, Peepal, Banyan, and Eucalyptus, along with seasonal sequestration potential of crops such as sugarcane, rice, wheat, and maize. I referred to afforestation and reforestation project data from various Indian states, and compiled land-based sequestration coefficients based on vegetation type, land use, and regional climate conditions. Additionally, I gathered information from companies working on carbon capture methods such as biochar, agroforestry, and soil carbon storage. The values were cross-verified using multiple references to ensure accuracy and contextual relevance.

5. Project Phase 2 – Scope 3 Data Categorization and Collection

5.1 Understanding Scope 3 Emission Categories

Scope 3 emissions represent indirect greenhouse gas emissions occurring across the **value chain** of an organization. Unlike Scope 1 (direct emissions) and Scope 2 (indirect emissions from purchased energy), Scope 3 encompasses **15 categories** that are often the largest portion of an entity's carbon footprint.

The categories I focused on during this internship included:

- **Category 1** – Purchased Goods and Services
- **Category 2** – Capital Goods
- **Category 3** – Fuel- and Energy-Related Activities
- **Category 4** – Upstream Transportation and Distribution
- **Category 8** – Upstream Leased Assets
- **Category 9** – Downstream Transportation and Distribution

Each category demanded specific data types, sources, and assumptions for accurate emission calculation.

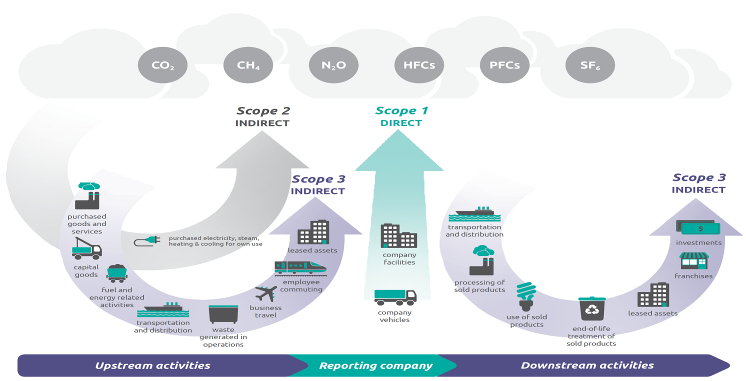

Figure 10: The 15 categories of Scope 3 GHG emissions across the value chain.

5.2 Scope of Work and Category-Wise Approach

Category 1: Purchased Goods and Services

- Focused on emissions from procurement activities (office supplies, IT hardware, services)
- Collected data on emission factors per unit cost. Used environmentally extended input-output (EEIO) databases to estimate emissions

Category 2: Capital Goods

- Analyzed emissions from **long-term assets** (machinery, electronics, construction)
- Used amortized emission factors per asset class
- Referred to data from Ecoinvent and capital goods reports

Category 3: Fuel- and Energy-Related Activities

- Captured upstream emissions from fuel extraction, processing, and transportation
- Collected data per unit of energy (kWh) and fuel type (diesel, coal, electricity grid mix)

Category 4: Upstream Transportation and Distribution

- Investigated emissions related to the **inbound movement of goods**
- Segregated by mode (road, rail, air) and distance
- Factored in vehicle efficiency and logistics load

Category 8: Upstream Leased Assets

- Looked at **assets used but not owned**, such as rental equipment or shared office space
- Emissions calculated based on usage hours and asset type
- Used average grid and asset-specific consumption rates

Category 9: Downstream Transportation and Distribution

- Focused on emissions from **product delivery** or distribution to end users
- Estimated using delivery distance, freight mode, and weight of goods
- Referred to DEFRA transport emission factors and logistics studies

5.3 Data Collection Sources

Source	Use Case
IPCC Guidelines	Emission factor standards
UK DEFRA Emission Factor Datasets	Transport & distribution
Ecoinvent Database (reports)	Capital goods, manufacturing
Indian Ministry of Power, BEE	Grid emission intensity, energy use
Business Sustainability Reports	Benchmark for leased assets and procurement data

5.4 Data Structuring and Formats

All collected data was structured into standardized **CSV and Excel files**, with the following fields:

- **Category Name**
- **Activity Description**
- **Emission Factor (kg COe/unit)**
- **Unit Type (, km, kWh, item, etc.)**
- **Source / Citation**

5.5 Python Scripts and Automation

Python scripts were created to:

- **Ingest raw data** from different formats
- **Normalize values** to a common base unit
- **Map activity data** to the correct Scope 3 category
- **Compute total emissions** by multiplying quantity × emission factor

6. Project Phase 3 – Life Cycle Assessment (LCA) using Brightway2 and OpenLCA

6.1 What is Life Cycle Assessment (LCA)?

Life Cycle Assessment (LCA) is a structured method for evaluating the environmental impacts of a product, process, or activity across its **entire life cycle** — from raw material extraction to disposal. It helps quantify emissions, energy usage, and resource consumption at each stage.

Stages in LCA:

- Raw material extraction
- Manufacturing
- Transportation
- Usage
- End-of-life disposal or recycling

6.2 Brightway2 Overview

Brightway2 is an open-source Python framework for LCA. It allows users to:

- Build LCA models programmatically
- Import and manage large datasets (e.g., ecoinvent)
- Perform custom environmental impact assessments
- Automate and customize entire workflows using code

It's ideal for **researchers, developers, and advanced users** comfortable with Python.

6.3 Brightway2 Workflow

The Brightway2 LCA workflow generally includes:

1. **Create/Select Project**
2. **Import Database (e.g., ecoinvent or custom)**
3. **Define Functional Unit** – quantity of product/service to assess
4. **Link Activities and Flows**

5. **Choose Impact Assessment Method** (e.g., IPCC, ReCiPe)
6. **Run LCA Calculation**
7. **Analyze Results**

6.4 Differences Between Brightway2 and OpenLCA

Aspect	Brightway2	OpenLCA
Interface	Code-based (Python)	GUI-based
Flexibility	High (custom models, automation)	Medium (limited scripting support)
Learning Curve	Steeper (requires coding)	Beginner-friendly
Dataset Integration	Ecoinvent, custom databases via Python	Direct support for multiple LCI formats
Ideal For	Developers, researchers	Students, analysts

6.5 Scope for Integration with Dashboard

LCA results from Brightway2 (or OpenLCA) can be exported as numerical values (e.g., kg COe) and then:

- **Integrated into the emissions dashboard** for lifecycle-level analysis
- **Used to compare suppliers or materials** used in event planning
- **Visualized** in Streamlit using charts and tables
- **Exported** in reports with per-person or per-item impact

7. Key Learnings

7.1 Technical Skills Gained

During the internship, I enhanced my skills in Python programming, data analysis, and backend logic implementation. I developed interactive dashboards using Streamlit, managed structured data with SQLite, and visualized results using Matplotlib and Plotly. I also worked with Brightway2 for life cycle assessment and applied data cleaning techniques using Pandas.

7.2 Domain Knowledge Acquired (Climate Science & Emissions)

I gained a solid understanding of Scope 1, 2, and 3 emissions, carbon sequestration methods, and the role of emission factors in sustainability reporting. The internship deepened my exposure to climate science and environmental modeling. and sustainable computing:

7.3 Communication and Team Collaboration

- Participated in regular progress check-ins and coordinated with mentors and peers.
- Maintained task documentation and version control to stay aligned with shared goals.
- Improved clarity in reporting technical and environmental data for both technical and non-technical audiences.

.

8. Conclusion

8.1 Overall Experience

The FOSSEE internship gave me a meaningful opportunity to **merge technology with climate action**. Through this project, I was able to go beyond basic development and explore how digital tools can contribute to real-world sustainability.

From understanding emissions to actually calculating and visualizing them, every step of the internship felt connected to a larger environmental goal.

8.2 What Worked Well

- Building a working **dashboard from scratch** using open tools
- Creating a **user-friendly form system** that functioned both online and on mobile
- Mapping real data to **emission factor models**
- Exploring the potential of **carbon sequestration** in future tools
- Independently managing code, research, and reporting

8.3 Scope for Future Work

- Add **user login system** and allow attendee profile editing
- Support **multiple simultaneous events** and event-wise reports
- Fetch emission factors from **real-time APIs**
- Use carbon sequestration offsets to calculate **net emissions**
- Expand dashboard to include **more Scope 3 categories**

9. References

Data Sources

1. Intergovernmental Panel on Climate Change (IPCC) Emission Factor Guidelines
2. Food and Agriculture Organization (FAO) – Global Emission Statistics
3. UK Department for Environment, Food & Rural Affairs (DEFRA) – GHG Conversion Factors
4. Government of India – Ministry of Road Transport and Highways (MoRTH) Reports
5. Indian Railways Passenger Emission Estimations
6. Ecoinvent Database Documentation (as applicable)
7. National Renewable Energy Laboratory (NREL) – LCA Data Sources
8. Brightway2 and OpenLCA Toolkits
9. Academic Journals on Carbon Sequestration (Elsevier, Springer)

Research Papers and Links

1. GHG Protocol Corporate Standard: https://ghgprotocol.org
2. IPCC Reports Archive: https://www.ipcc.ch/reports
3. Brightway2 Documentation: https://brightway.dev
4. DEFRA Emission Factors (UK Gov): https://www.gov.uk/government/collections/government-conversion-factors-for-company-reporting
5. FAO Climate Smart Agriculture: https://www.fao.org/climate-smart-agriculture/en/
6. OpenLCA Software: https://www.openlca.org
7. FOSSEE Internship Portal: https://fossee.in