Sri Lanka Institute of Information Technology

PROJECT REGISTRATION FORM

(This form should be completed and uploaded to the Cloud space on or before XXXXXXXXX)

The purpose of this form is to allow final year students of the B.Sc. (Hon) degree program to enlist in the final year project group. Enlisting in a project entails specifying the project title and the details of four members in the group, the internal supervisor (compulsory), external supervisor (may be from the industry) and indicating a brief description of the project. The description of the project entered on this form will not be considered as the formal project proposal. It should however indicate the scope of the project and provide the main potential outcome.

PROJECT TITLE (As per the accepted topic assessment form)	Real-time Carbon Neutrality Management and Optimization using Natural Language Processing
RESEARCH GROUP **(as per the Topic assessment Form)**	Knowledge Inspired Computing (KIC), Data Science (DS)
PROJECT NUMBER	**TMP-22-081** (will be assigned by the lecture in charge)

PROJECT GROUP MEMBER DETAILS: (Please start with group leader's details)

	STUDENT NAME	STUDENT NO.	CONTACT NO.	EMAIL ADDRESS
Format	Perera C.D.D	ITxxxxxxxx	0712345678	itxxxxxxxx@my.sliit.lk
1	Sathees P. (GROUP LEADER)	IT19052748	0777609352	it19052748@my.sliit.lk
2	Mathanika M	IT19005218	0766351696	it19005218@my.sliit.lk
3	Vishakanan S	IT19001562	0768306757	it19001562@my.sliit.lk
4	Vithursan M.	IT19033174	0772630659	it19033174@my.sliit.lk

SUPERVISOR, CO_ SUPERVISOR Details

SUPERVISOR Name	**CO-SUPERVISOR Name**
Ms. Anjali Gamage	Ms. Sanjeevi Chandrasiri
Signature	**Signature**
Attach the email as Appendix 1	**Attach the email as Appendix 2**
Date	**Date**

EXTERNAL SUPERVISOR Details (if any, may be from the industry)

Dr. Daniel N Subramaniam	Experienced in managing carbon emissions for various organizations.	-	0212060160	**Attach the email as Appendix 3**
Name	Affiliation	Contact Address	Contact Numbers	Signature/Date

ACCEPTANCE BY CDAP MEMBER (This part will be filled by the RP team)

Name	Signature	Date

PROJECT DETAILS

Brief Description of your Research Problem: (extract from the topic assessment form)

In realizing the Paris agreement to lessen the consequences of climate change, legislation for a drastic reduction of greenhouse gas (GHG) emissions was imposed by governments worldwide [1]. As a part of this measure, organizations must disclose their GHG emissions publicly [2]. Some governments implement strategies such as carbon credit (cap on total GHG emissions that businesses are allowed to emit. E.g., one carbon credit = 1 metric ton of GHG emission), carbon offsetting (businesses pay to have the emission reduced somewhere else. E.g., Solar power projects), and carbon in setting (investing in emission reduction within the business supply chain. E.g., In premise solar power supply). Businesses go "carbon neutral" by keeping the total emissions within the cap, and businesses go "carbon neutral." This reporting task is the organization's responsibility, usually handled by a business analyst (BA). For this, all emission tasks (e.g., use of 5l of gasoline for an electrical generator) must be recorded, and emission values should be calculated by comparing the relevant emission factor (e.g., 1l of gasoline will emit 0.25 kg of CO2 equivalent) [3]. Governments will release the data sets of these emissions factors each year (e.g., DEFRA standard by UK government) [4]. This task is overwhelming and requires a thorough understanding of this process. Moreover, there is no way of knowing the current emission compared to their target (emission cap). Therefore, there is a need for real-time monitoring and optimization of the organizations' GHG emissions [5].

References

[1] 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: https://www.nature.com/articles/s41467-020-15414-6?fbclid=IwAR1drArL9ReoJl2zgqjmdxJNoBsM4zRJna-JHIGWkzTka7d4NB4fdz0nCrE.

[2] S. Tang and D. Demeritt, "Climate Change and Mandatory Carbon Reporting: Impacts on Business Process and Performance", *Business Strategy and the Environment*, vol. 27, no. 4, pp. 437-455, 2017. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/bse.1985.

[3] M. Brander, M. Gillenwater and F. Ascui, "Creative accounting: A critical perspective on the market-based method for reporting purchased electricity (scope 2) emissions", *Energy Policy*, vol. 112, pp. 29-33, 2018. Available: https://www.sciencedirect.com/science/article/pii/S0301421517306213.

[4] "Measuring and reporting environmental impacts: guidance for businesses", *GOV.UK*, 2021. [Online]. Available: https://www.gov.uk/guidance/measuring-and-reporting-environmental-impacts-guidance-for-businesses. [Accessed: 16- Dec- 2021].

[5] 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.

Sub problem 1 – Efficient emission factor searching

Most of the countries' governments that part takes in the Paris agreement [1] release quite an extensive emission factor document for various activities, annually [2]. These factor values are used to calculate final emission values for the activities performed over the reporting period [2]. It is not so easy to find relevant emission factors as these are mentioned in technical terms and require a deeper understanding of them. Even though there are several commercially available emission calculators available, those also require users to have in-depth knowledge about the emission factors and want the users to select relevant emission factors [3]. As emission factors must be found for each activity, reducing the effort and time involved here would result in increased productivity for carbon reporting.

References

[1] 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: https://www.nature.com/articles/s41467-020-15414-6?fbclid=IwAR1drArL9ReoJl2zgqjmdxJNoBsM4zRJna-JHIGWkzTka7d4NB4fdz0nCrE. [Accessed 16 December 2021].

[2] "Measuring and reporting environmental impacts: guidance for businesses", *GOV.UK*, 2021. [Online]. Available: https://www.gov.uk/guidance/measuring-and-reporting-environmental-impacts-guidance-for-businesses. [Accessed: 16- Dec- 2021].

[3] S. Tang and D. Demeritt, "Climate Change and Mandatory Carbon Reporting: Impacts on Business Process and Performance", *Business Strategy and the Environment*, vol. 27, no. 4, pp. 437-455, 2017. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/bse.1985. [Accessed 16 December 2021].

Sub problem 2 – Real-time collection of emission data

It is usually the responsibility of a single business analyst (BA) or a small BA team to maintain the emission activities of their organization, calculate emissions and produce reports. Collecting data on emission activities from various sources is a tedious task for a business analyst and might be erroneous on some occasions [1]. Since this is a time-consuming task and only performed only a few times a month, there will be no real-time up-to-date status of the organization's emissions. This is also true for most of the current online emission calculators since they also require users to take the data collection responsibility.

References

[1] Bebbington, J. and Larrinaga-González, C., 2008. Carbon Trading: Accounting and Reporting Issues. *European Accounting Review*, [online] 17(4), pp.697-717. Available at: https://www.tandfonline.com/doi/full/10.1080/09638180802489162 [Accessed 23 December 2021].

Sub problem 3 – Unit checking and conversion to match the emission factor's unit

In the emission details collection using a voice, assistant users are not limited to entering certain units and emission parts extraction would not be aware of the different units it only knows to classify it as a unit. Emission factor search will also result in most matching and not limited to units and users can select emission factors with different units. Therefore, before calculating emission values, user-given values' units should be verified with selected emission factors unit and must be converted before calculating emission [1].

References

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

Sub problem 4 – Emission optimization for the emission source constraints

Even though reducing carbon emission is a crucial task every organization must carry out, there are occasions in which it is still not possible to consider alternative sustainable options yet. Most of these emission sources are related directly to the business process and there will be no alternative option that will not affect the business effectiveness. However, there are many emission sources in an organization that can be reduced to achieve their desired emission goal [1]. There should be a way to find an optimal solution for these kinds of constraints on emission sources [2].

References

[1] Ibrahim, M., Putri, M. and Utama, D., 2020. A literature review on reducing carbon emission from supply chain system: drivers, barriers, performance indicators, and practices. *IOP Conference Series: Materials Science and Engineering*, [online] 722(1), p.012034. Available at: https://iopscience.iop.org/article/10.1088/1757-899X/722/1/012034/meta [Accessed 23 December 2021].

[2] Sarkar, B., Omair, M. and Choi, S., 2018. A Multi-Objective Optimization of Energy, Economic, and Carbon Emission in a Production Model under Sustainable Supply Chain Management. *Applied Sciences*, [online] 8(10), p.1744. Available at: https://www.mdpi.com/2076-3417/8/10/1744 [Accessed 23 December 2021].

Description of the Solution: (extract from the topic assessment form)

An innovative solution proposed for the above scenario would be to implement a real-time platform that can provide insights on the most up-to-date emission statistics of the organization. Emission activity data will be directly gathered from the employees using a **voice assistant**. For each emission activity input, relevant emission factors will be found with the help of an **information retrieval** process. Before proceeding, any misinformation will be clarified with the employee. To avoid miscalculations, units of inputs will be checked by using **text classification** with **natural language processing** and values will be converted before emission calculation. calculated emission values will be stored for analysis purposes and business analysts can access this real-time data using any business intelligence tool. For the constraints provided on the emission sources and emission cap, an optimal solution would be generated by using **optimization models (algorithms)** and alerts will be sent if there is any breach of the optimal solution.

Sub solution 1 - Efficient emission factor searching

Using the emission activity parts identified by the voice assistant, a factor searching module will provide relevant factors back to the voice assistant for confirmation. Since the emission factors are provided in a technical language, a **natural language-based information retrieval** system will be used to conveniently find relevant emission factors. It will find the relevancy of a factor to the query, based on closeness, term frequency, and personalization weightage. This approach can be scaled to many different emission factor standards.

Sub solution 2 - Real-time collection of emission data using a voice assistant

For faster input of emission tasks, a **voice assistant** will be implemented. Employees can provide emission records using natural languages. Necessary parts will be obtained from the audio file using **natural language processing** approaches. In the event of ambiguity or missing parts, the voice assistant will request the user to clarify. Moreover, it will get confirmation on emission factors found as relevant.

Sub solution 3 - Unit checking and conversion to match the emission factor's units

Classes of the unit provided by the user and unit in the selected emission factor will be classified using a **text classification** approach using **natural language processing**. Using the identified classed similarity of the units will be verified if the classes are the same. If those classes are different, the conversion factor for unit change will be derived using a **conversion matrix** for different unit classes and values will be converter using the conversion factor before calculating emission.

Sub solution 4 - Emission optimization for the emission source constraints

Details on usage constraints on several emission sources will be provided by the business analyst. For these constraints and emission cap, a minimal optimal solution will be provided using **optimization models**. This can be customized by the business analysts according to their requirement changes. Once an optimal solution is confirmed, the maximum emission value provided by the solution for various emission sources will be set as the max thresholds and if these values are exceeded alerts will be sent.

Main expected outcomes of the project: (extract from the topic assessment form)

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

Sub Objective 1:
Search emission factors and provide ranked results based on closeness, term frequency, and personalization weightage.

Sub Objective 2:
Gather employee emission activity details from employees using a voice assistant.

Sub Objective 3:
Verify and convert values for units provided by the employees to match the units of the selected emission factor.

Sub Objective 4:
Identify the optimum solution for the given emission source constraints and alert about any violations of the optimal solution.

WORKLOAD ALLOCATION (**extract from the topic assessment form after the correction suggested by the topic assessment panel.**)
(Please provide a brief description about the workload allocation)

MEMBER 1	Sathees P. - IT19052748

1. Collecting emission standard data from the emission standard publication websites.
   e.g., DEFRA (Department for Environment, Food and Rural Affairs) – UK, CRIS (Climate Registry Information System) – USA and Canada, EPA (Environmental Protection Agency) – USA, NGA (National Greenhouse Accounts) – Australia
2. Implementing an emission factor searching module that can provide **ranked results** for the query based on different factors such as **closeness, term frequency, personalization weightage**.
   - Using emission activity and unit as the inputs, this component will return the ranked results of matching emission factors for the employee to choose.
     a. Perform necessary **decoding, preprocessing, cleaning, and transformation** on the acquired emission factor dataset files.
     b. Identify a common **document structure** to store different emission standards.
     c. Process the cleaned dataset into the common document structure and store it in document storage with associated fields such as **year, standard name, scope, emission values, units, and corpus (factor fields)**.
     d. Implement a suitable **indexing algorithm** and with the help of it create and store the index after indexing the corpus.

e. Implement custom **IR (Information Retrieval) models** with support for ranking based on the search **closeness, term frequency, and personalization weightage**. This model returns **ranked results** in reverse order (Most matching in the top) with the maximum confidence metric of each search.
f. Using the previous user search and selection history as feedback, implement a **personalization** feature to the IR model.

3. Implement **emission calculation** for the emission activity and emission factor selected by the user and store it for further analysis.
    a. Calculate emission from user-fed activity data such as **amount, unit, time, and emission conversion factor chosen**.
    b. Store these values in a relational database for application purposes.
    c. Additionally, store the emission data in a data warehouse (in a Dimensional model) from the application to facilitate real-time analysis.

MEMBER 2 | Mathanika M - IT19005218

1. Make new training data for emission parts extraction model by conducting online surveys to different organization employees. Manually label the necessary parts from the provided texts.
2. Implement a **voice assistant** to gather emission activities from the employees in real-time.
    - Using the emission tasks provided by the employees in an audio format, this component will extract parts necessary for emission calculation.
    a. Extract voice data in an audio format from the employees using the mobile app.
    b. Implement a **compression system** on the mobile device to compress the captured voice using a compression algorithm.
    c. Implement a **conversion model** to change voice data into text.
    d. Train a **natural language processing model** to extract necessary emission parts (values, units, type of emission, time) from the input text.
    e. Check the validity of the extracted parts and if the validity criteria are not met, more clarification will be obtained from the employee using the voice assistant again for the missing part.
    f. Provide extracted parts emission factor search module and retrieve results with a maximum confidence metric. If the confidence metric is too low, provide comments on having low confidence.
    g. On confirmation of the emission factor value, provide the parts and selected emission factor value for emission calculation.
    h. Ask the user to choose text input if the noise is too high.
3. Provide analysis access to the organization's emission data for business analyst usage.
    a. Segment data in the data warehouse according to the organization
    b. Provide separate access to the segmented data via an off-the-shelf business intelligent tool.

MEMBER 3 Vishakanan S - IT19001562

1. Make new datasets of different classes of units and different representations of them from an online survey and manually label the representation.
2. Implementing text classification model with natural language processing to verify the difference in input units and units in the emission factor.
   - For a given unit in the emission activity and an emission factor, will identify whether these units are the same.
     a. Train a **text classification model** to identify classes of the units. This should be tolerant of different unit representations.
     b. Identify classes of both input unit and emission factor's unit.
     c. Verify whether these classes are the same. If they are the same send the values to emission calculation.
3. Creating a unit conversion system if the provided units' classes are different.
   - Values provided will be converted to the units of emission factor's unit if the units are different.
     a. Create a **unit conversion factor matrix** for all the unit classes.
     b. Identify the unit conversion factor for the given unit classes from the implemented unit conversion factor matrix.
     c. Convert values to the emission factor's unit using the identified unit conversion factor and send it to emission calculation.

MEMBER 4 Vithursan M. - IT19033174

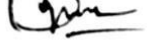
1. Implementing a custom emission optimization module for the given constraints on different sources.
   - Using the constraints on emission sources as the input, this component will return an optimal solution to minimize the emission.
     a. Obtain usage constraints on different emission sources of the organization from the business analyst (BA).
     b. Implement an **optimization model** using an appropriate **optimization algorithm** to find the minimum solution for the given emission sources constraints and total emission cap of the organization.
     c. Let BA configure and choose a suitable optimal solution.
     d. Create thresholds on different emission sources according to the chosen optimal solution.
2. Creating an alert framework to provide alerts about the breaches of the thresholds added by the optimal solution.
     a. During the addition of a new emission record, check whether any thresholds provided by the optimal solution are violated.
     b. Send alerts to the BAs about any threshold violations.

3. Implement a mobile application using React Native and expo cli. React Native will be used to implement a cross-platform mobile application and expo cli will be used to access get hardware components such as mic and speaker.

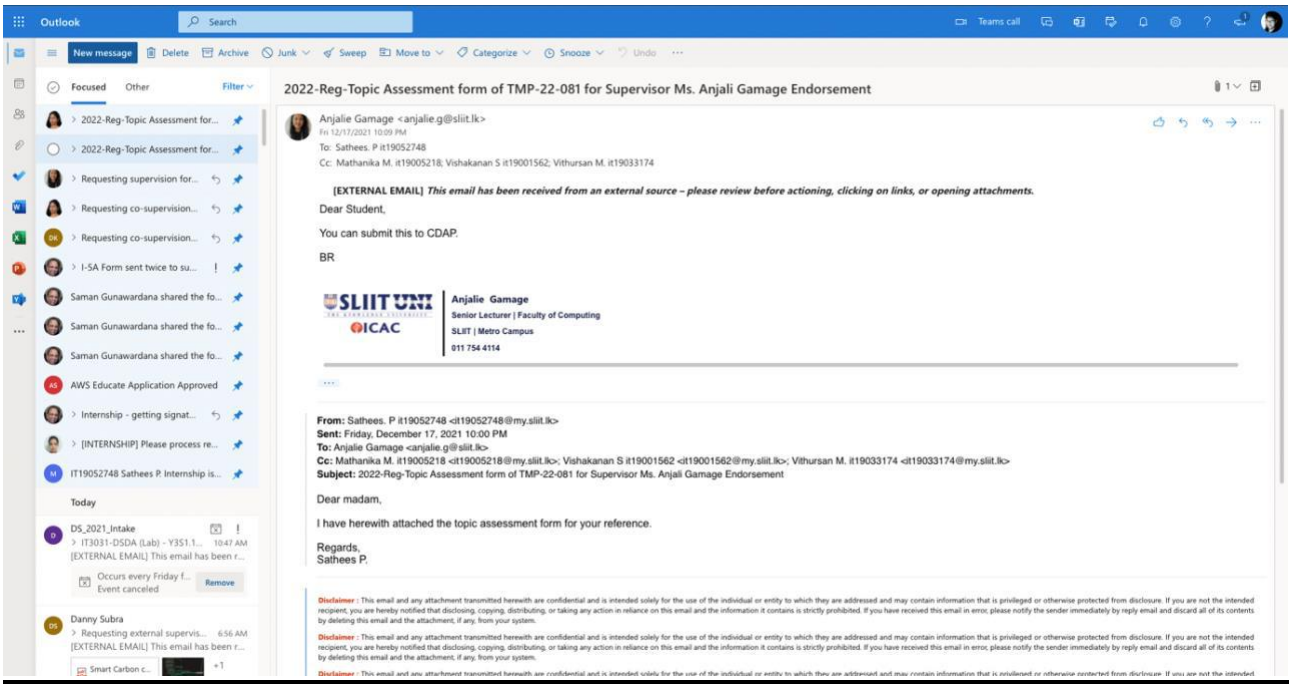
DECLARATION (Students should add the Digital Signature)

"We declare that the project would involve material prepared by the Group members and that it would not fully or partially incorporate any material prepared by other persons for a fee or free of charge or that it would include material previously submitted by a candidate for a Degree or Diploma in any other University or Institute of Higher Learning and that, to the best of our knowledge and belief, it would not incorporate any material previously published or written by another person in relation to another project except with prior written approval from the supervisor and/or the coordinator of such project and that such unauthorized reproductions will construe offences punishable under the SLIIT Regulations.

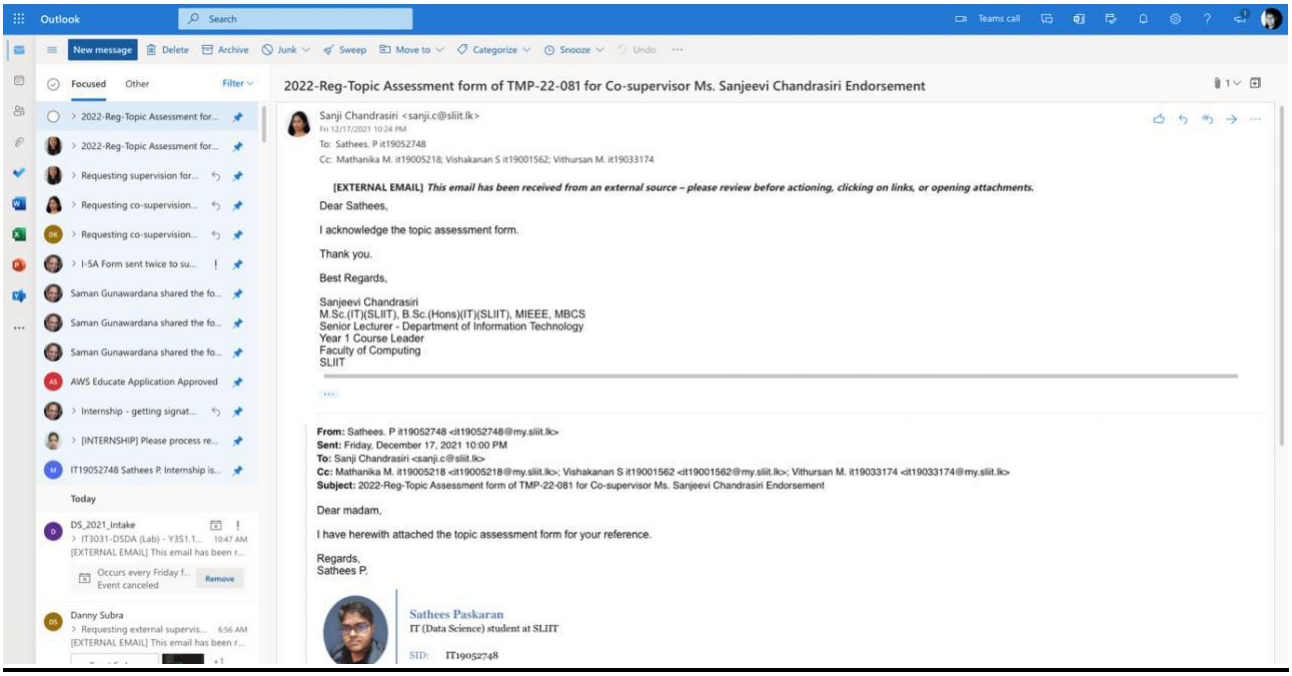
We are aware, that if we are found guilty for the above mentioned offences or any project related plagiarism, the SLIIT has right to suspend the project at any time and or to suspend us from the examination and or from the Institution for minimum period of one year".

	STUDENT NAME	STUDENT NO.	SIGNATURE
1	Sathees P. (GROUP LEADER)	IT19052748	
2	Mathanika M	IT19005218	
3	Vishakanan S	IT19001562	
4	Vithursan M.	IT19033174	

Appendix 1 – Supervisor e-mail endorsement



Appendix 2 – Co-supervisor e-mail endorsement



Appendix 3 – External supervisor e-mail endorsement