

THE GREEN TRANSFORMATION LAB

The Singapore Management University is collaborating with DHL, the world's leading logistics provider to accelerate the evolution of sustainable logistics across Asia Pacific through the Green Transformation Lab.



IS482 Project Proposal

Prepared for:
Professor Seema Chokshi

Professor Tan Kar Way
Ms Michelle Kan
Mr Andrew Beng

Janice Koh Rui Ling | Tan Siong Min | Tay Hui Shia

Background

In the society today, the pace of globalization is undeniable. With remarkable increase in transportation, communication and technology, it has made the world more interdependent than ever. However, as companies reap in the success of globalization, Mother Earth sacrifices.

Resource depletion and severe environmental issues has not only motivated the increasing interest towards renewable resources but also, sustainability policies. The limelight on sustainability development has proved to be a valuable area for development reaping huge benefits. Businesses are now moving to take an active role in developing practices promoting sustainability with initiatives such as sustainable logistics.

The Green Transformation Lab (GTL) is a joint initiative by SMU and DHL aimed at accelerating the evolution of sustainable logistics across Asia Pacific. Leveraging SMU's multi-faculty academic excellence and DHL's sustainability services, expertise and capability in supply chains, the Green Transformation Lab focuses on creating solutions that help companies transform their supply chains, becoming greener, more resource efficient and sustainable.

We will be embarking on gaining further insights on GTL's sustainability heat map project, which is a group of heat maps with global profiling of areas and trends on sustainability related jobs as well as sustainability related topics on Twitter.

Motivation

Due to the increasing gap between companies' sustainability initiatives and the workforce involvement, more efforts are placed in integrating sustainability concepts into business models and jobs so as to create a greater impact on the environment.

Through this project, we are interested to understand the elements powering the creation of sustainability jobs, be it government policies, workforce preference or the pervasive impact of social media.

Objectives

The main objective of our project is to analyze crawled real-time tweets as well as job information related to sustainability and environment at a deeper level so as to gain more insights, relationships and trends. We will also be exploring various environmental and socio-economic indicators data justified by the United Nations (UN). Through multiple analytical methodologies, we hope to answer the below project objectives.

- 1) From the crawled real-time tweets from Twitter, explore the relationship between highly related topics and impact generated by posts that are labeled with the hash tag "#sustainability".
 - a) Which topics has the highest frequency of being mentioned with "#sustainability"?
 - b) Define the type and genre of users posting tweets related to sustainability (based on level of popularity and influence of the users)
 - c) Are companies generating tweets related to sustainability? Differentiate the topics driven by social networks and commercial companies.
 - d) Are users whose tweets contain sustainability-related hash tags, connected to one another (following/follower relationship)? Are there any cascade effects on the sustainability tweets?
- 2) To determine the correlation between real-time tweets and job information related to sustainability and environment.
 - a) What are the impacts of this relationship?
 - b) Are there more jobs created in the respective categories due to this correlation?
- 3) Based on the crawled data on sustainability jobs, is there an association of policies (e.g. country policies, governmental policies) to sustainability jobs?
 - a) What are the effects on the creation of sustainability jobs due to government policies such as sustainable development and resource efficiency?
 - b) Are nations' environmental watchdogs contributing to the spike in sustainability jobs in the recent years?
- 4) To find out whether other environmental and socio-economic indicators have a form of impact sustainability jobs based on locations such as Europe, Asia and America.
 - a) Identify the effects of various macro-economic metrics and UN environmental indicators to determine whether there is an existing trend with sustainability jobs in the various regions.

Scope

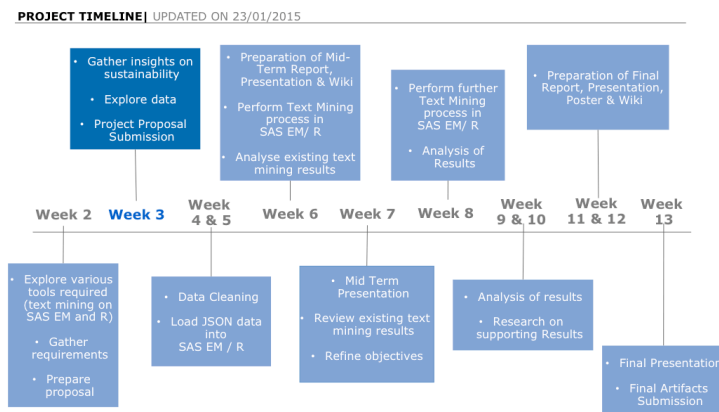
We will be focused on text mining to extract meaningful indices of the crawled data, both sustainability job and twitter, thereby inputting into predictive and data modeling techniques. With text analysis, it would allow the discovery of new information and topics that will aid us in meeting our project objectives. We would also explore appropriate global environmental and socio-economical indicators through the UN database and extracting various dataset. Using multiple analytical methodologies, we seek to provide better insights into the existing dataset provided to us.

Recommended Solutions

We will be exploring software like SAS Enterprise Miner and R to gain insights on the crawled sustainability data provided from the job searches and Twitter. Through Enterprise Miner, we will embark on descriptive and predictive modeling, thereby developing models that provide insights on understanding the job information relating to sustainability and environment.

In addition, text mining could also be run in R using the **tm** package. This **tm** package consists methods of data import, corpus handling, preprocessing, metadata management and creation of term-document matrices¹. We will start with loading the corpus (collection of text), which in this case is our crawled data, exploring and preparing the corpus, creating and documenting the text matrix, identifying the frequent items and association and thus creating the correlation plots. Based on the results and through performing statistical methodologies, we hope to answer the questions proposed in our objectives.

Project Timeline



Limitations & Assumptions

In order to specify the scope our project, here are our existing assumptions and limitations:

- Assume information on crawled real-time tweets are accurate and may have a direct influence on job market.
- The aggregation of the yearly environmental and socio-economic indicators data would probably affect the accuracy level of the results when compared to the daily crawled sustainability jobs and Twitter data.
- The lack of location tagging on tweets might limit the analysis of the relationship between sustainability-related tweets and sustainability jobs by regions.

Possible Avenues for extension of the scope of the project

Integration may be done with the current sustainability heat map website² to showcase the analysis and trends of our project research. This information would then be publicly available for users to understand and gain insights of sustainability jobs around the world and how it is impacting the world today.

Acknowledgements and Credit

Sponsor: Professor Tan Kar Way, Michelle Kan (IT Solutions Project Manager at GTL SMU) and Andrew Beng (IT Solution Engineer at GTL DHL)

Supervisor: Professor Seema Chokshi

¹ Feinerer. Introduction to the tm package. Text Mining in R. June 2014. URL: <http://cran.r-project.org/web/packages/tm/vignettes/tm.pdf>

² <http://heatmap.greentransformationlab.com>