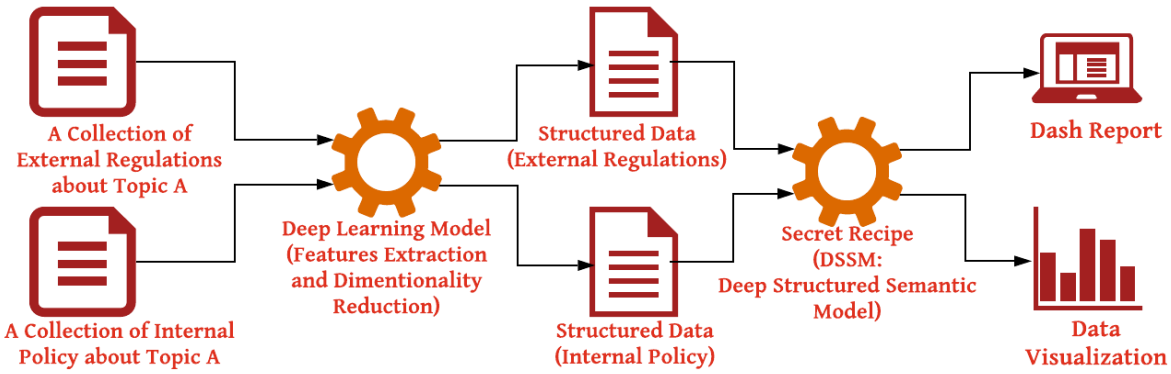
Quantify Your Risk Model (QYR Model)

---- By Shawn Hu, Vicky Zhou, Wenting Zhao

Quantity Your Risk (QYR) offers a quantified solution powered by Deep Structured Semantic Learning Method, NLP, machine learning, dashboard metrics, and data visualization to close the gap between external regulations and internal controls. The model offers many benefits as it not only minimizes the risk of internal control failure but also significantly reduces the cost and labor time of Citi's compliance team. QYR can immediately spot external regulations' **risk and cost level. Through** quantifying the risk of asymmetrical internal controls and external policies, it generates actionable control insights **to reduce the crisis response time** and **costs** of internal control failure. It **automates** Citi's compliance processes significantly **reducing compliance labor time and costs**. QYR offers a **self-improving and robust** solution that helps Citi **constantly monitor** their risk and improve their overall control environment.

Quantity Your Risk Model (QYR Model) :



1. Feature engineering & Intent Modeling (Deep Learning NLP): The "**Deep Learning Model**" has embedded **glossary lists** about each policy domain. The model **extracts key features** from raw policy documents. Then the model transforms **unstructured raw documents** into **numerical matrices** that can **describe the meaning and intent** of each policy based on various statistical algorithms. Each row of the matrices represents one policy document and each column stands for one key feature. Finally, it implements **dimensionality reduction algorithms** to select **meaningful key features** for **dash board data visualization**.

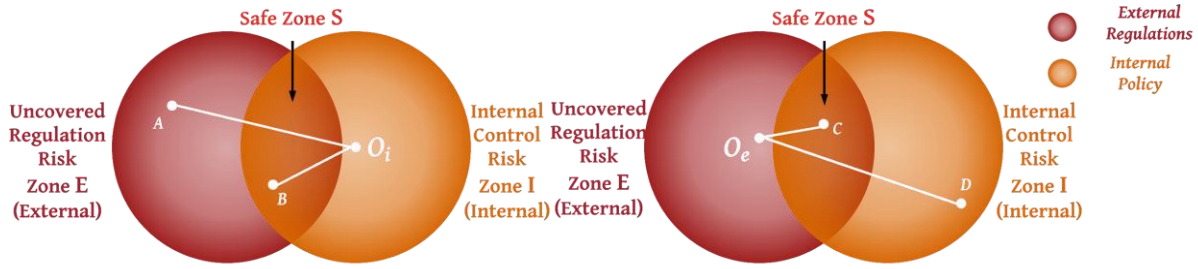
2. Risk Assessment: **Secret Recipe (DSSM) [1]** takes the two **numerical matrices** generated from step 1 and then identify the location of each policy document geometrically. The location of each regulation or internal control indicates its corresponding risk level. The "**Deep Structured Semantic Model**" translates the **gap** between internal and external policies into numerical **Risk-Cost Metrics**, which can **quantify** the **overall risk of the control environment** as well as **the risk and cost associated with a specific external regulation or internal control.** Risk Metrics are applicable **to all kinds of regulation updates, drafts,** and **designed controls**.

3. Standardization and Optimization**:** The solution **immediately generates dashboard reports and informs** the compliance department high-risk segments of controls, regulations, and entities. After the staff feeds the newly designed controls back into **QYR**, it recomputes Risk Metrics **and evaluates the overall risk of the control environment until it reaches Citi's standard and meets the requirements of the regulations.**

4.Constant Monitor: The model **restarts from step 1** updating new draft regulations and internal controls into RYS; it evaluates Risk-Cost Metrics constantly.

Secret Recipe (Deep Structured Semantic Algorithm DSSM) [1] is the key component of the QYR model, which is shown below.

[1] https://posenhuang.github.io/papers/cikm2013_DSSM_fullversion.pdf

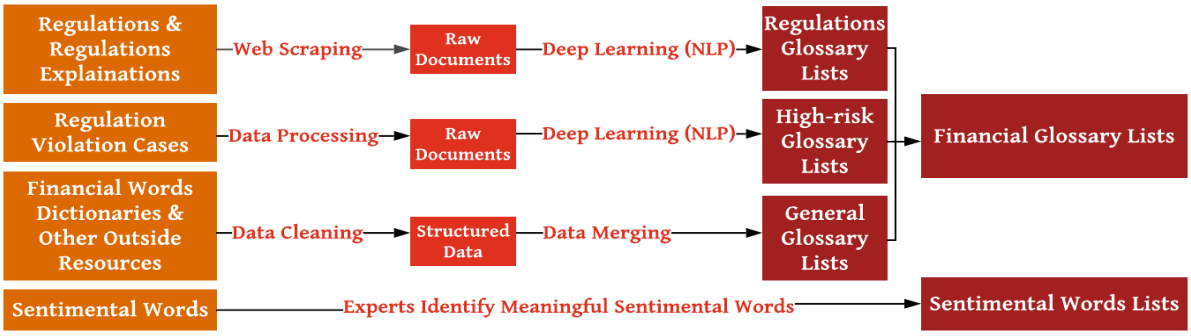


Secret Recipe (DSSM Algorithm): These two graphs are **a simplification of DSSM Model.** After feature engineering (step 1), DSSM inputs a matrix of external regulations (E) and a matrix of internal policies (I) within the same domain and **calculate the Risk Metrics to identify the overall risk level and assign risk indexes to all the internal controls and external regulations**.

Risk Metrics refer are **meaningful numbers** that can be applied to reduce risk and costs. A variety of Risk-Cost Metrics will be provided by the model to provide more risk insights. Each white point in the graph above stands for an external regulation or an internal policy. The distance between the white point and the center of the opposite circle measures the **distance**(**risk**) of mismatch internal control and external regulation. However, a gap does not mean **costs for Citi**. The model multiplies a cost level of a specific regulation, predefined by compliance experts, and the distance, between the policy and its opposite circle center, to calculate **its cost level**.

Two graphs diagnosis: The left graph captures the **risk-cost level** when Citi fails to comply with **a specific external regulation**. The right graph captures the **risk-cost level** caused by **specific internal control**. For example**,** Red **Circle** and **Orange Circle** represent a cluster of related external regulations and internal controls regarding the conduct of **Broker-Dealer (BD)**. A new external regulation **A** regarding **BD comes out and** immediately feeds into **QYR. B represents a safe spot** as it lies within the **overlapping region (Safe Zone). Clearly, the distance A-Oi is greater than B-Oi, which implies regulation A is at higher risk to violate than B and needed immediate attention.**

To build QYR model, we **can start with the "Deep Learning Model**" in the first graph. The first step is to create the embedded glossary lists(**features**) and train the "Deep Learning Model". Other key features will also be included but not shown below.



The graph above represents the processes to generate embedded glossary lists and create the "deep learning model" for **Feature Engineering** in which we identify **all the key features that describe the intent of a specific external regulation or internal policy.** The steps are as follows:

1. The Model Collect regulations from the official websites of SEC, FINRA and CFTC; It transforms the raw documents into a list of regulation glossaries based on word frequency using NLP.
2. Collect all regulation violation cases and transform the unstructured text into a list of words related to high risk glossaries.
3. Process and merge all financial terminologies defined by compliance experts into the glossary lists.
4. Collect sentimental words predefined by compliance experts and transform them into a list of words.

Sample Code: https://github.com/wenting94/PwC_case