SMART INDIA HACKATHON 2019

Technology Bucket : **Inconsistency Detection in Medical Annotation**

Category: **Software**

Company Name: **ezDI**

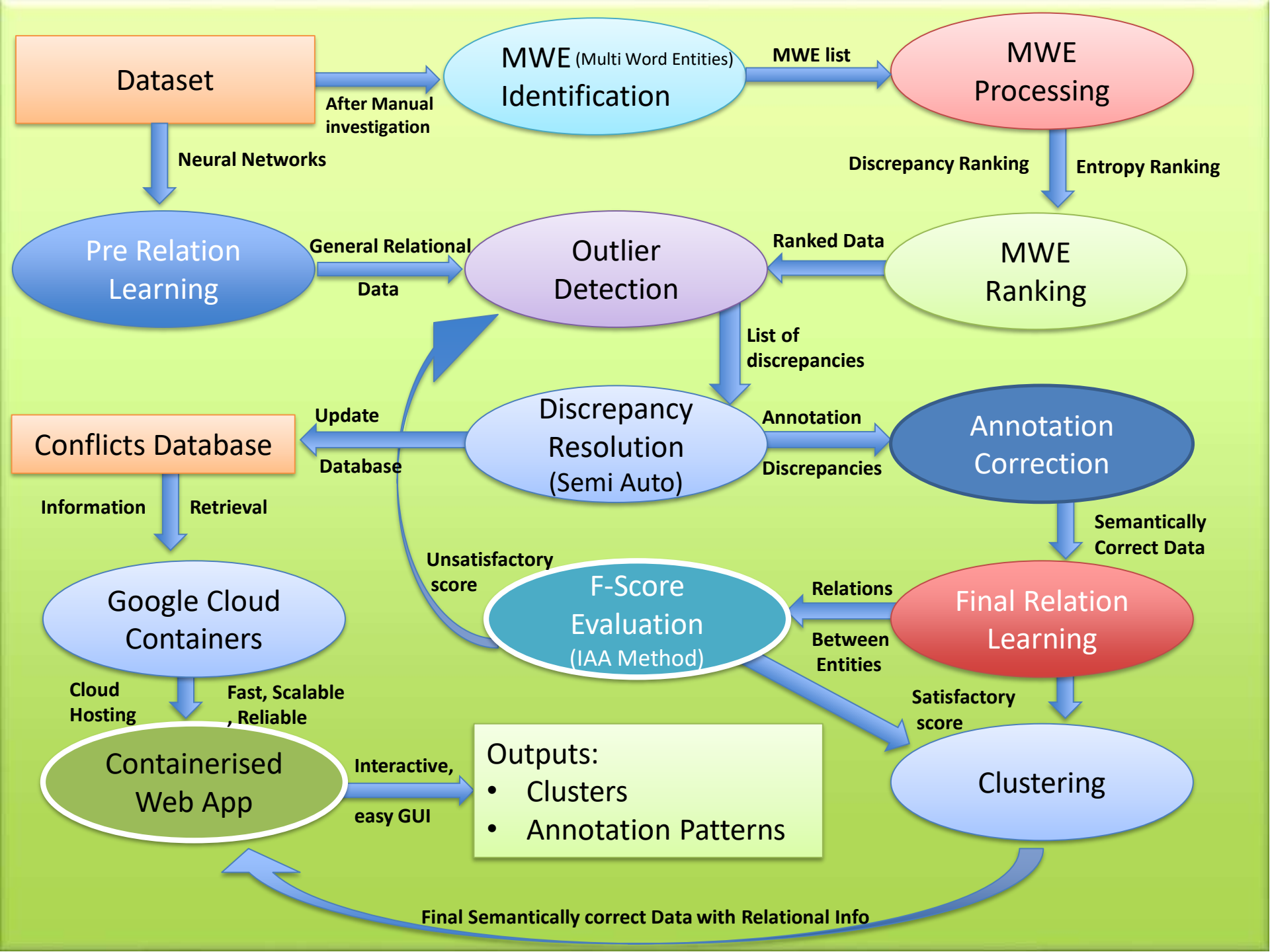
Problem Code : **RG1**

Team Leader Name : **Amrit Preet Singh**

College Code : **1-3514263743**

Solution in brief:

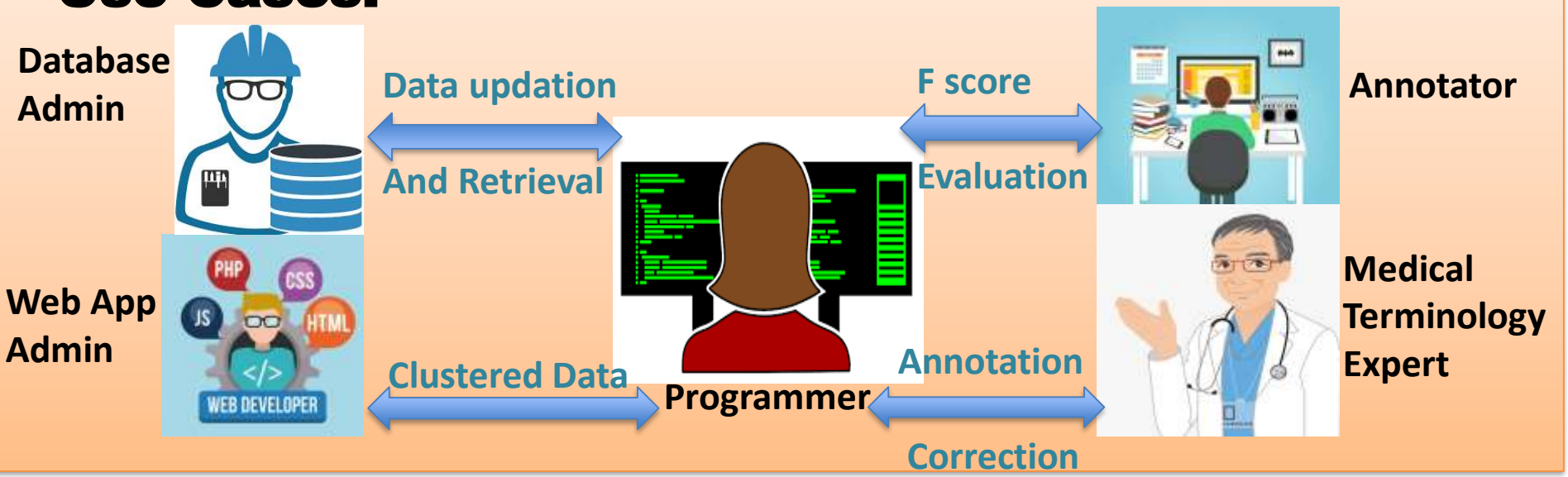
We propose a dynamic system for clustering the similar entities and to associate the annotation patterns which will be updated in a databse maintained using MySQL which will be fed to a web application hosted on Google Cloud which will display the result in a fast, scalable and reliable manner. The detailed architecture of the system is shown below.


Dataset
After Manual investigation
MWE (Multi Word Entities) Identification
MWE list
MWE Processing
Neural Networks
Discrepancy Ranking
Entropy Ranking
Pre Relation Learning
General Relational Data
Outlier Detection
Ranked Data
MWE Ranking
List of discrepancies
Update Database
Discrepancy Resolution (Semi Auto)
Annotation Discrepancies
Annotation Correction
Conflicts Database
Information Retrieval
Semantically Correct Data
Unsatisfactory score
F-Score Evaluation (IAA Method)
Relations Between Entities
Final Relation Learning
Google Cloud Containers
Cloud Hosting
Fast, Scalable , Reliable
Satisfactory score
Containerised Web App
Interactive, easy GUI
Outputs:
• Clusters
• Annotation Patterns
Clustering
Final Semantically correct Data with Relational Info

Technology Stack:

- Tensorflow, Keras, Neural Networks, Tensorboard (for visualisation).
- Machine Learning Algorithms for Clustering data points.
- Python 3.6, Java, HTML, CSS, Javascript.
- Google Cloud.
- MySQL

Use Cases:



Dependencies:

1. Dataset which ezDI will provide.
2. Atleast 6 GB GPUs for training of our model(we'll arrange it if not provided).

Future Work: We propose to use this web app to build a Tensorflow Lite app which will automate the process of annotating new rtf data and automatically perform clustering and relationship mappings