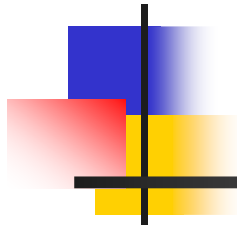


Interlingua on the Web:

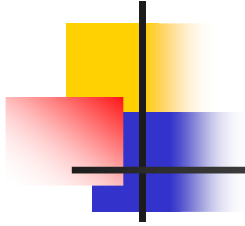
AgroExplorer

Guide: Prof. Pushpak Bhattacharyya



Sarvjeet Singh

Mrugank Surve



Goals:

- **Efficiency:** What is the time required for the retrieval of Documents
- **Scalability:** Can the system handle web sized Corpus
- **Accuracy:** What is the Precision and Recall of the system



Drawbacks of the Earlier System

- **Inefficient:** did not Index Documents
- **Not Scalable:** could not handle a large corpus
- **Used Complete Matching:** missed some relevant Documents



Improvements:

- **Indexing:**

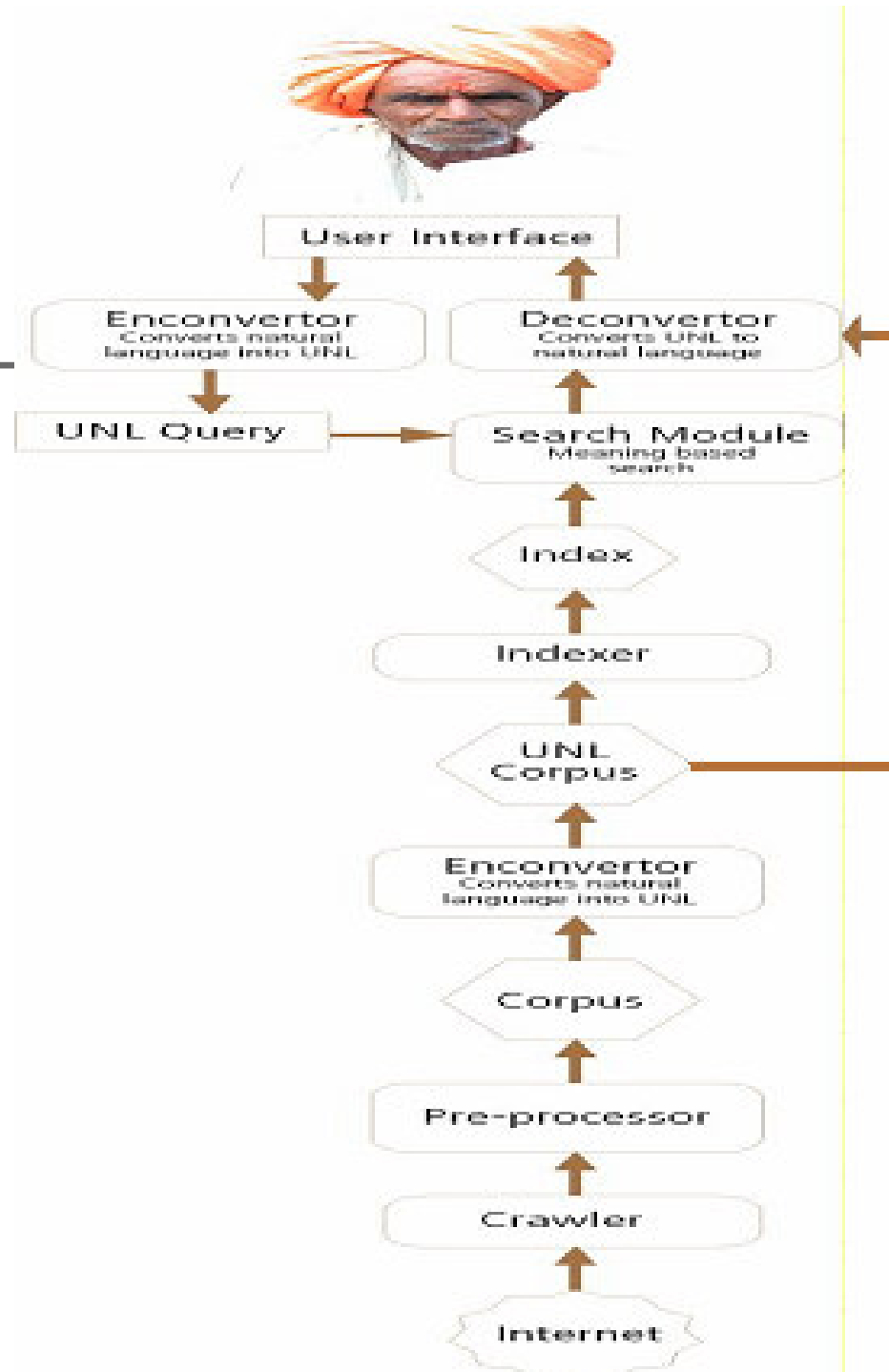
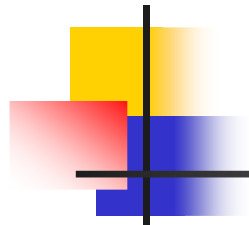
- Index Documents based on the UNL expressions

- **Partial Matching:**

- Also Retrieve Documents having a partial match with the query

- **Graph Based Search:**

- Searches for the query graph in the sentence graph





Milestones Achieved:

- **Indexing Module**
- **Search Module**
- **Database Interface to Search Module**



Future Work:

- **Focused Crawler**
- **HTML Parser**
- **User Interface**
- **Integration of the Various Modules**