

# Character Identification and Classification



Team Members: Geeticka Chauhan, Labiba Jahan (VIP volunteer)

Product Owner: Mark Finlayson

Instructor: Masoud Sadjadi, Francisco Ortega

School of Computing and Information Sciences

Florida International University

# Goal

To understand stories; science for cultural understanding

Tool:  
Natural  
Language  
Processing

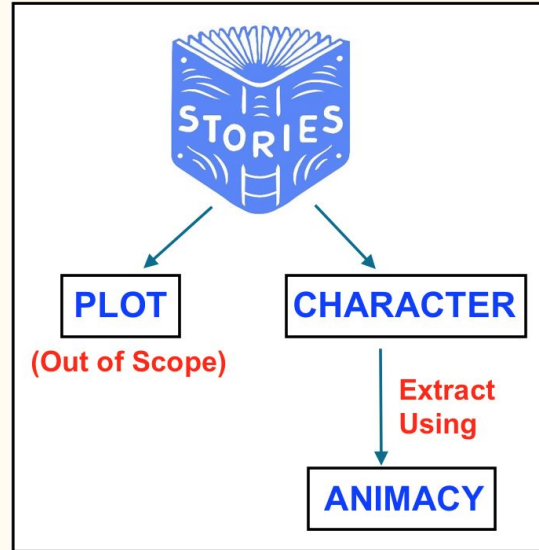


Image Source: slate.com and IBM Watson

Geeticka Chauhan

# Terminology

## Referring Expressions, Coreference Chains and Animacy

Rebecca Ross has been studying for a test, and she has been using flash cards to help her out.

He used the stove, and it was hot.

The tree said, “Hi, my name is Steve.”

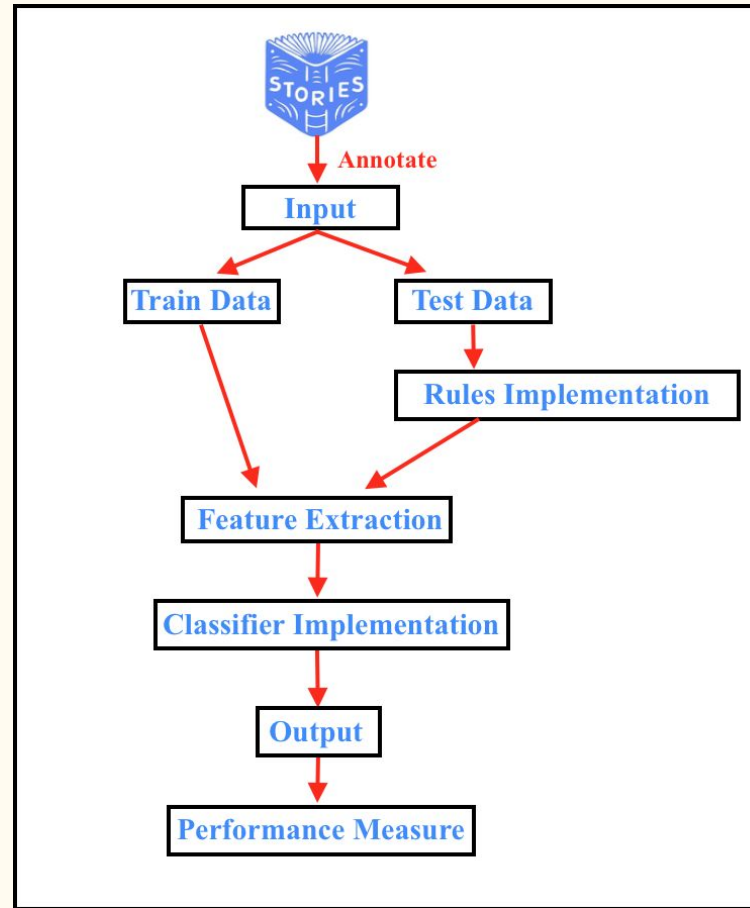
**animate** **inanimate** ← Stove --- it

Steve --- The tree

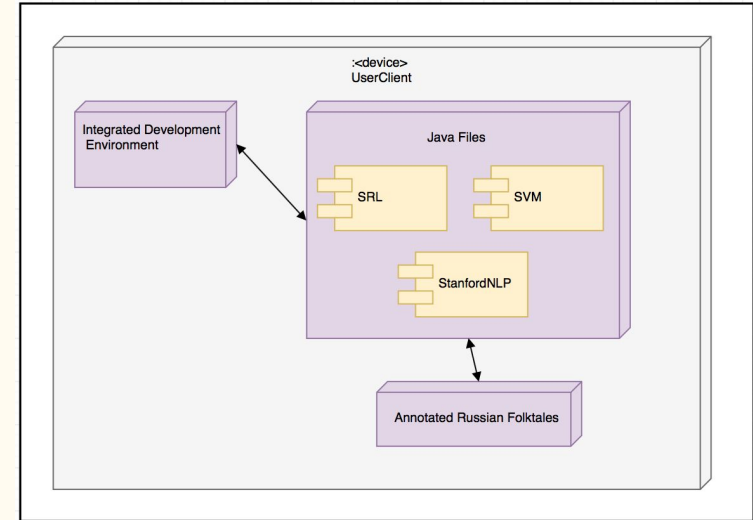
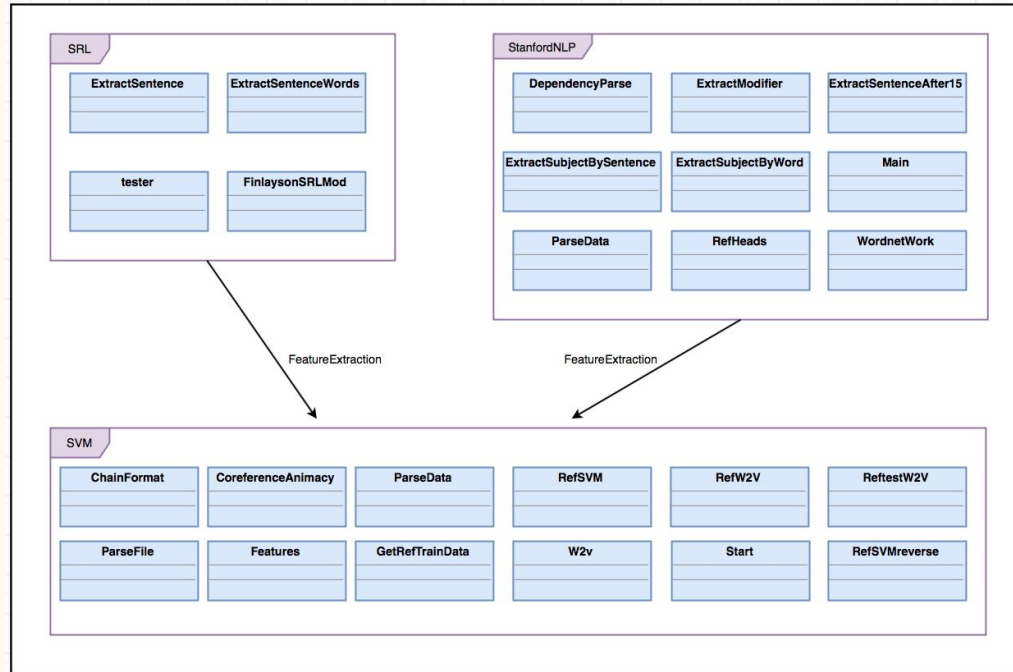
# Process

**Input:** Annotated  
Russian Folktales

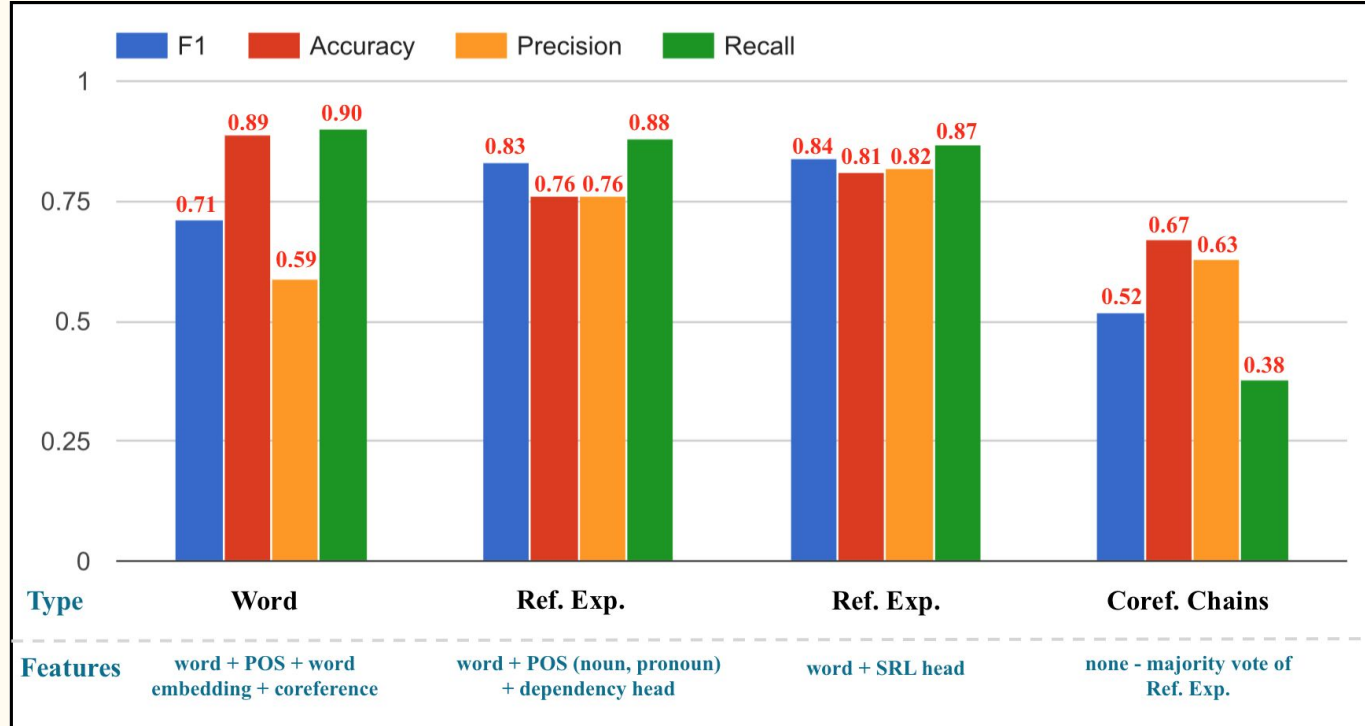
**Output:** Animacy of  
Coreference Chains



# Subsystems and Deployment



# Results



# Contributions

1. Created clean data
2. Experimental Setup of Animacy Classification
3. Devised SVM Classifier
4. Good results of 84% F1

## Future Work

### Ongoing

**Rule-Based System - improve result of Coref Chains**

### Next Steps

**Automatically extracting character name + roles**

## Technologies Used

Stanford NLP  
MIT Java WordNet Interface  
MIT Story Workbench  
Word2Vec Java Library



## Contact Info

**GCHAU005@fiu.edu**  
**LJAH002@fiu.edu**  
**MARKAF@fiu.edu**