Finding Trolls Under Bridges:
Preliminary Work on a Motif Detector

W. Victor H. Yarlott[1], Armando Ochoa[1], Anurag Acharya[1], Laurel Bobrow[2],
Diego Castro Estrada[1], Diana Gomez[1], Joan Zheng[2], David McDonald[2],
Chris Miller[2], Mark A. Finlayson[1]

[1]Knight Foundation School of Computing and Information Sciences, FIU
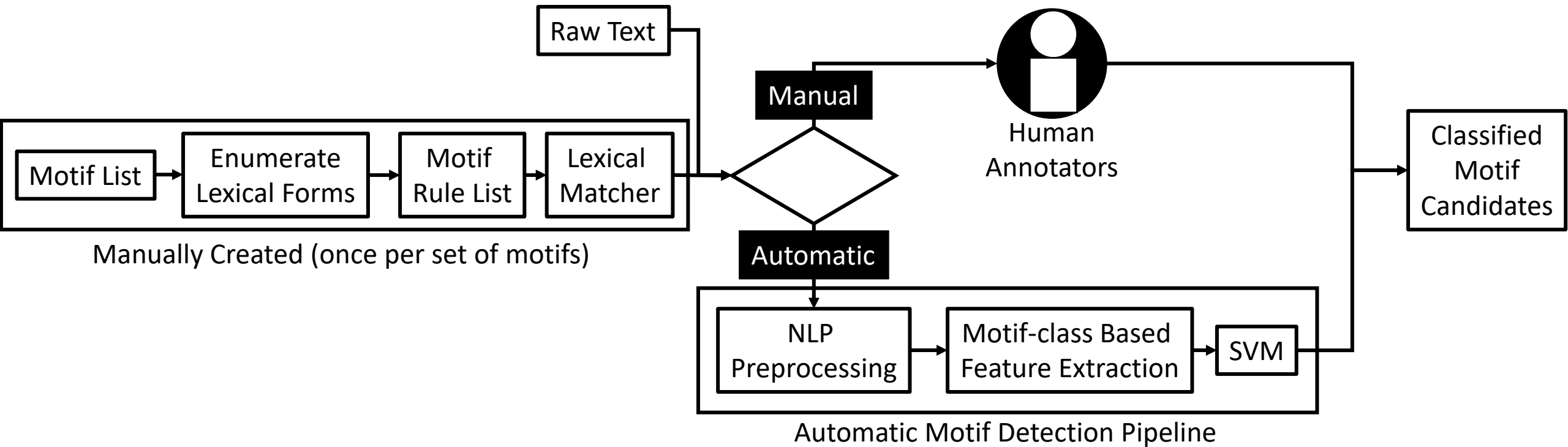[2]Smart Information Flow Technologies (SIFT)

Motivations

- What *are* motifs?
  - Distinct, recurring narrative elements found in folklore and, more generally, culturally inflected materials
  - Can be generally divided into character, prop, and event
  - For example, the character "troll under the bridge" →

- Motifs are ubiquitous
- Embody a wealth of related ideas

Motif Annotation Pipeline

Raw Text

Manual

Human Annotators

Motif List → Enumerate Lexical Forms → Motif Rule List → Lexical Matcher

Manually Created (once per set of motifs)

Automatic

Classified Motif Candidates

NLP Preprocessing → Motif-class Based Feature Extraction → SVM

Automatic Motif Detection Pipeline

Data Production & Human Annotation

- Motif selection
    1. Identifying cultural group to draw motifs:
        1. Irish, Jewish, and Puerto Rican
    2. Finding highly-referenced motifs
    3. Finding definitive source within folklore
    4. Estimate the strength of motifs

- Double annotation is done by three pairs of annotators

- Annotate candidates as motific, referential, eponymic, or unrelated

- Annotators are given training with weekly adjudication sessions

- Annotators use brat and are given a motif annotation guide

- Have reached agreement of Fleiss' kappa > 0.55

Irish, Finn McCool
  a young **Finn McCool** with limbs that might bestride a bullock
Jewish, Tower of Babel
  communism had crumbled like the **Tower of Babel**
Puerto Rican, Coqui
  We listen here for the first strains of the exotic, in the parrot's squawk, the **coqui's** clamor, the reggaeton beat

Architecture

- Preprocessing
  - Standard NLP + targeted preprocessing

- Feature Extraction
  - Character motifs: animacy, coreference
  - Prop motifs: semantic, grammatical, possessional relations
  - Event motifs: event detectors, semantic roles

- Classified using SVM

Preliminary Results: Metaphor

Model	Motif F_1	Eponym F_1	Referential F_1	Unrelated F_1	Macro F_1
Metaphor	0.35	0.00	0.59	0.00	0.21

- Some overlap between motif and metaphor

- Metaphor better at detecting referential usage

- Relation between metaphor and motif is not 1-to-1

- Referential usage may use more literal language

Future Work & Contributions

- Future Work
  - Continued development of motif detector
  - Completion of annotation study
- Contributions
  - Shown preliminary agreement measure suggesting humans with the relevant background can reliably distinguish motific usage of motif surface forms
  - Presented in-development architecture for automatic motif detection
  - Presented preliminary results on the use of metaphoricity as a feature for distinguishing motific usage of motif surface forms

Slack Channel: #paper23-yarlott