

Foreign Language Tool Improvement Through Evaluation (FLITE)

Flo Reeder

703-983-7156 • freeder@mitre.org

MITRE Sponsored Research

The logo for the MITRE Technology Program, featuring a stylized graphic of stacked blocks in yellow, orange, and blue to the left of the text.

MITRE
Technology
Program

The MITRE logo, consisting of the word "MITRE" in a bold, black, sans-serif font.

MITRE

© 2004, The MITRE Corporation

Problem

- **Primarily government-funded companies for languages of interest – limited and slow**
 - **Cleared, qualified linguists not available**
- **MT is a hard problem, but if constrain the domain & expectations, it is a useful tool**
 - **Rapid prototype solutions**
 - **TIDES' surprise language experiment**
 - **Pushing evaluation**
 - **Full automation of evaluation**
 - **Look at psycholinguistics research**
 - **Why is a translation good?**

Background

- **What is a good translation?**
 - **No single right answer to that question**
 - **Steer MT evaluation to task-based metrics**
 - **Steer users to better understanding of what MT can do for them**
 - **Push to automate metrics that correlate well with judgments of quality**
- **Improve existing tools through automated evaluation – use machine learning techniques**
- **Give natural language generation attention – often ignored before**

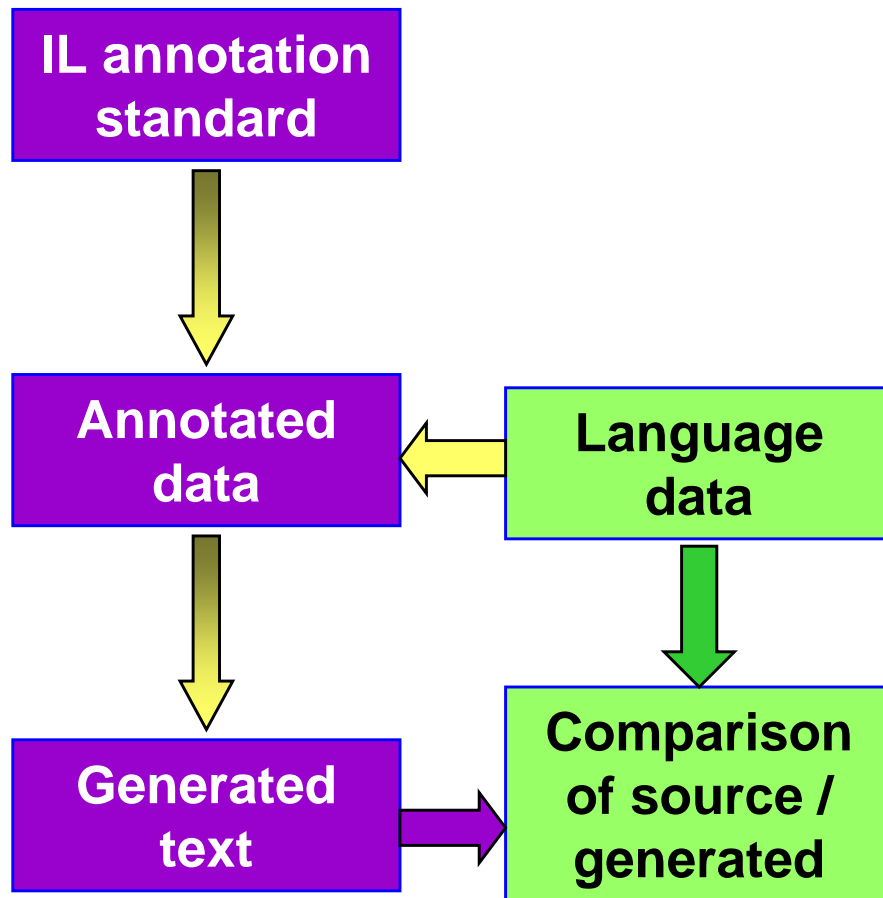
Objective

- **Improve language tools by incorporating evaluation into the processing**
- **Pushing the boundaries:**
 - **Transformation-Based Learning of MT – Building MT with small data sets**
 - **Multi-engine evaluation platform – MuMMS**
 - **Lexicon learning from comparable corpora – fixed-point translation focus - ELLIS**
 - **Improved natural language generation – generate from Inter-Lingual Form**

Activities

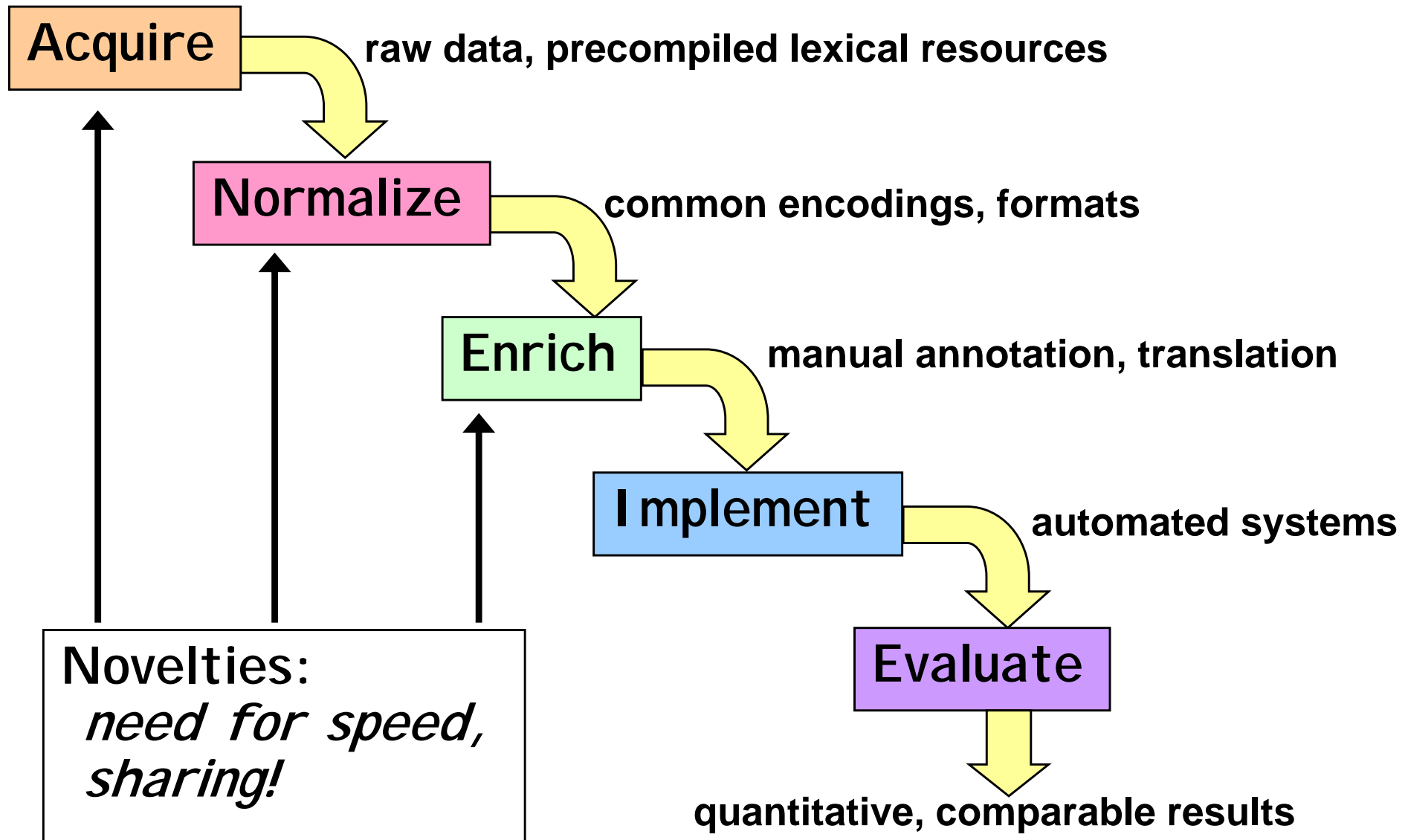
- Transformation-based learning – Incorporation of BLEU metric into TBL program, adaptation of BLEU metric
- Development of Ellis – fixed-point term translation learning tool
- Multi-engine, multi-metric evaluation framework (MuMMS) – transition to sponsor
- Development of non-reference based metrics
- Surprise Language Experiment
- Development of interlingual representation

Highlight: Getting Knowledge from Language



- Breaking semantic wall requires understanding language and how it works
- Annotating data and learning from this is core element in process
- Design interlingua for MT which:
 - Uses available semantic information – ontologies
 - Expands to new languages
 - Can be evaluated

Highlight: Building MT Components



Impacts

- **Development of new metrics for MT**
 - Paris – n-graph engine tested by NIST
 - Nidaba – suite of individual metrics
- **Ellis – fixed-point translation finder**
- **Multi-engine, multi-metric evaluation system (MuMMS)**
- **TBL iterations ongoing**
- **Use of Halogen generator with IL Format**
- **Investigation of psycholinguistic research**

Future Plans

- **Transformation-based learning**
 - **Bootstrapping with fixed-point translations**
 - **Look at learning part-of-speech tagging, parsing strategies**
 - **Segmenting data into morphemes**
- **Evaluate without reference translations**
- **Breaking the sentence boundary**
- **Strive for interlingua, semantics, context**
- **Begin looking at new communication modes**